

UNIVERSIDAD DE BURGOS
PROGRAMA INTERNACIONAL DE DOCTORADO
ENSEÑANZA DE LAS CIENCIAS

Departamento de Didácticas Específicas



**TESTES DE HIPÓTESES: ANÁLISE DO DOMÍNIO
DESTE CAMPO CONCEITUAL POR ALUNOS DA ÁREA
DA SAÚDE**

TESIS DOCTORAL

Rodrigo Fioravanti Pereira

Burgos, mayo de 2020

UNIVERSIDAD DE BURGOS
PROGRAMA INTERNACIONAL DE DOCTORADO
ENSEÑANZA DE LAS CIENCIAS

Departamento de Didácticas Específicas



UNIVERSIDAD DE BURGOS



**UNIVERSIDADE FEDERAL
DO RIO GRANDE DO SUL**

Tesis Doctoral

**TESTES DE HIPÓTESES: ANÁLISE DO DOMÍNIO
DESTE CAMPO CONCEITUAL POR ALUNOS DA ÁREA
DA SAÚDE**

Tese de D. Rodrigo Fioravanti Pereira apresentada à
Escola de Doctorado da Universidad de Burgos, para a
obtenção de Título de Doutor em Ensino de Ciências

Orientadores: Dra. Ileana Greca Dufranc y Dr. Jesus
Meneses Villagra

Burgos, 29 de mayo de 2020

Rodrigo Fioravanti Pereira

TESTES DE HIPÓTESES: ANÁLISE DO DOMÍNIO
DESTE CAMPO CONCEITUAL POR ALUNOS DA ÁREA
DA SAÚDE

290 páginas

Tese - Escuela de Doctorado de la Universidad de
Burgos.

1. Didáctica da Bioestatística

2. Área da Saúde

3. Testes de Hipóteses

I. Universidad de Burgos. Escuela de Doctorado.
Departamento de Didácticas Específicas.

Dedicatória

Letícia, Mariana, Ana, Maguins, Sirlei e Deroci.

Conseguimos!

Não deveríamos avaliar as ações humanas com base nos resultados.

Bernoulli

Agradecimentos

Amo meu trabalho! Adorei esta pesquisa!

Me orgulho de cada vírgula neste texto, ele reflete a minha verdade. É claro que não foi escrito sozinho e é claro que não vou conseguir agradecer a todas as mãos que guiaram a minha, mas não vou deixar de tentar. Lá vai...

A profe Ileana é uma das pessoas mais capazes que eu conheço. Ela e o Profe Meneses pairam em ares rarefeitos demais pra mim. Aprendi muito sob a orientação deles e com a onipresença do Profe Moreira. Muito Obrigado!

Mas aprendizagem também é troca. E foi trocando ideias com os meus amigos e companheiros Letícia Fogaça, Letícia Stefenon e Léo Dalla Porta que descobrimos muito do que Vergnaud tinha a dizer. Muito Obrigado!

A troca de ideias sobre educação sempre foi comum pra mim, minha mãe, Dona Ana Marília, fez história como Professora e o meu pai, Seu Maguins, tem colecionado talentos ao longo da vida. Eles são uma fonte de inspiração, possas exemplares com as quais todos deveriam conviver. Muito Obrigado!

Jornadas que valem a pena não são solitárias! Eu tive a melhor das companhias, minha esposa Letícia Fogaça. Fui Professor e Aluno dela. Ela me descifra e me completa, me eleva e me liberta e ainda fez o filho único ter mais dois pais, Seu Dero e Dona Leca. Muito Obrigado Meu Amor!

Eu queria ser perfeito por causa Dela. Um pedaço meu que vive fora e dentro de mim. Espelho no qual me reconheço. Amor incondicional. Orgulho incomparável. Mari, o pai te ama mais que a ele mesmo! Muito Obrigado!

Palavras tentam concretizar sentimentos, por isto é que agradecimentos escritos são sempre insuficientes. Não posso nominar aqui todos aqueles que levo no coração, a todos estes... Muito Obrigado! Saibam que valeu a pena!

Sério, eu queria dizer que foi problemático, difícil, ..., aquela coisa toda, mas olha só quem esteve comigo... Foi bom demais!

Resumo

A prática baseada em evidências é importante para a área da saúde, e a estatística, um de seus pilares, é presença comum nos currículos da graduação. A capacidade computacional e a quantidade de dados disponíveis, popularizaram o uso da estatística na pesquisa desta área. Contudo, não se observa um número significativo de pesquisas sobre a didática da estatística neste contexto. A revisão de literatura identificou, ainda, que o ensino de estatística sofre com alunos propensos à ansiedade estatística, pouco uso das teorias de aprendizagem e que há falta de instrumentos e técnicas de análise quantitativas para o estabelecimento do nível de aprendizado dos estudantes desta área. Com vistas a contribuir com o ensino de estatística para a área da saúde, esta pesquisa propôs uma metodologia baseada no *Problem Based Learning* (PBL), apoiada na Teoria dos Campos Conceituais (TCC) a fim de avaliar em que medida se estabeleceu o domínio do campo conceitual (CC) dos Testes de Hipóteses para a Média Populacional (THMP) após sua aplicação a alunos de uma universidade do sul do Brasil e comparou os resultados com outros que experimentaram o ensino tradicional, na mesma instituição. A metodologia de pesquisa utilizou da triangulação das abordagens qualitativas e quantitativas. A análise qualitativa descritiva foi sobre os dados do Diário de Classe (DC) e das respostas a um instrumento aberto constituído de duas questões típicas dos THMP, propostas ao final das atividades. A abordagem quantitativa analisou este último instrumento levando em conta uma tabela de níveis alcançados em categorias capazes de indicar indícios de invariantes operatórios necessários para a resolução adequadas das questões. Houve, ainda, a comparação do rendimento dos estudantes, utilizando um instrumento fechado, com dez questões típicas dos THMP, cuja análise exploratória dos dados precedeu os testes estatísticos de comparação de médias para dois grupos independentes, conduzidos com os softwares Excel, Tableau, Orange, SPSS

e R. O estudo piloto revelou uma dificuldade do PBL em se adequar ao tempo exíguo do currículo, incitando a geração de uma estruturação do conteúdo para agilizar o processo autônomo de domínio do CC dos THMP. Também percebeu-se a necessidade da criação da ferramenta pedagógica chamada Oportunidade Didática (OD) a qual, pela ação do professor, e à luz da TCC, fomentou o domínio do CC no estudante por meio de perguntas as quais visaram a desacomodação ou consolidação dos aparentes invariantes operatórios disponibilizados em situação. Neste contexto, estabeleceu-se o ambiente de pesquisa apregoado na revisão de literatura. Os resultados apontaram a média de pontuação dos alunos participantes da metodologia levemente menor do que a dos não participantes, assim como no nível de domínio do CC dos THMP. As impressões coletadas no DC indicam que os participantes da metodologia se envolveram mais ativamente nas aulas e nos trabalhos em grupo, trazendo abordagens do conteúdo por vezes distinta das apresentadas no formato tradicional. Avalia-se que a metodologia de pesquisa foi adequada à análise dos dados, permitindo a análise do rendimento e do domínio do CC dos THMP. A metodologia de ensino mostrou-se factível mesmo num ambiente tradicional de ensino, onde os alunos apresentaram ansiedade estatística e baixas aspirações quantitativas, o professor contou com pouco auxílio dos colegas professores da área da saúde e o tempo de aplicação foi bastante curto. Mesmo neste cenário adverso, adaptado ao ensino tradicional, houve resultados não significativamente menores em relação ao tradicional, com o ganho de participação e atividade propositiva dos alunos durante as aulas e nos trabalhos em grupo, configurando um aluno mais propenso à pesquisa e ao desenvolvimento autônomo e participativo em relação ao conteúdo dos THMP.

Palavras-chave: Ensino de Estatística; Área da Saúde; *Problem Based Learning*.

Abstract

Evidence-based practice is important for the health area, and statistics, one of its pillars, is a common presence in undergraduate curricula. The computational capacity and the amount of available data popularized the use of statistics in research in this area. However, there is not a significant number of studies on the teaching of statistics in this context. The literature review also identified that the teaching of statistics suffers from students prone to statistical anxiety, little use of learning theories and that there is a lack of instruments and techniques for quantitative analysis to establish the level of learning of students in this area. To contribute to the teaching of statistics for the health area, this research proposed a methodology based on textit Problem Based Learning (PBL), supported by the Theory of Conceptual Fields (TCC) to assess the extent to which established the domain of the conceptual field (CC) of the Hypothesis Tests for the Population Average (THMP) after its application to students of a university in the south of Brazil and compared the results with others who tried traditional teaching, in the same institution. The research methodology used the triangulation of qualitative and quantitative approaches. The descriptive qualitative analysis was about the data from the Class Diary (DC) and the responses to an open instrument consisting of two typical THMP questions, proposed at the end of the activities. The quantitative approach analyzed this last instrument taking into account a table of levels reached in categories capable of indicating indications of operative invariants necessary for the adequate resolution of the questions. There was also a comparison of the students' performance, using a closed instrument, with ten typical THMP questions, whose exploratory data analysis preceded the statistical tests of comparison of means for two independent groups, conducted with the Excel, Tableau, Orange software. , SPSS and R. The pilot study revealed a difficulty for the PBL in adapting to the short

time of the curriculum, encouraging the generation of the structuring of the content to streamline the autonomous process of controlling the THMP CC. It was also noticed the need to create a pedagogical tool called Didactic Opportunity (OD) which, through the action of the teacher, and in the light of the TCC, fostered the student's control over the CC through questions which aimed at the discomfort or consolidation of the students. apparent operative invariants available in situation. In this context, the research environment proclaimed in the literature review was established. The results showed the average score of students participating in the methodology slightly lower than that of non-participants, as well as in the domain level of the THMP CC. The impressions collected in DC indicate that the participants in the methodology became more actively involved in classes and group work, bringing approaches to the content sometimes different from those presented in the traditional format. It is evaluated that the research methodology was adequate for the data analysis, allowing the analysis of the performance and the domain of the THMP CC. The teaching methodology proved to be feasible even in a traditional teaching environment, where students showed statistical anxiety and low quantitative aspirations, the teacher had little help from fellow health professors and the application time was quite short. Even in this adverse scenario, adapted to traditional teaching, there were results not significantly lower than the traditional one, with the gain of participation and purposeful activity of students during classes and in group work, configuring a student more prone to research and autonomous development and participatory about the THMP content.

Keywords: Statistics Teaching; Health Area; *Problem Based Learning*.

Lista de Figuras

1.1	Exemplo de saída dada pelo <i>software</i> SPSS, com mais de uma abordagem.	3
2.1	Estrutura metodológica da revisão própria de literatura	20
3.1	Relação entre os enfoques psicológicos utilizados nesta tese.	43
3.2	Imagem veiculada na matéria “Taxa de mortalidade por coronavírus no Reino Unido: o coronavírus matou mais do que a SARS - Qual é o perigo?” (Kettley, 2020)	57
3.3	Parte de um exame densitometria óssea digital de uma paciente de 83 anos de idade.	59
3.4	Resultado do exame densitometria óssea digital da 3.3.	60
3.5	Relação entre literacia (LITERACY), raciocínio (REASONING) e pensamento (THINKING) estatístico dada por Delmas (2002, p. 5)	63
3.6	Relação entre População, Amostra e Inferência.	68
3.7	Localização dos THMP dentro da estrutura da Estatística.	71
3.8	Regiões de rejeição de H_0 para cada tipo de H_1	74
3.9	Exemplo de p-valor menor do que α	78
3.10	Mapa conceitual do campo conceitual que envolvem o Teste de Hipóteses para a Média Populacional (THMP).	79
3.11	Constituintes do CC dos Testes de Hipóteses.	80

3.12	Ideia mínima esperada que o estudante tenha sobre o processo do THMP.	85
3.13	Zona de rejeição para uma distribuição Z e $\alpha = 5\%$.	88
3.14	Opções de caminhos a seguir para o desenvolvimento dos Testes de Hipóteses.	92
3.15	Exemplo de saída do <i>software</i> estatístico Bioestat.	93
3.16	Exemplo de saída do <i>software</i> estatístico Bioestat.	93
3.17	Representação da localização da estatística do teste frente a região crítica.	94
3.18	Etapas do <i>Problem Based Learning</i> .	98
4.1	Desenho da metodologia de pesquisa.	102
4.2	Desenho explicativo do instrumento para avaliar o nível de domínio do CC dos THMP.	105
4.3	Desenho explicativo do instrumento para avaliar o rendimento dos alunos acerca dos THMP.	106
4.4	Medidas resumo, por turma, das distâncias do semestre no qual o aluno fez a disciplina de Bioestatística em relação ao semestre indicado pelo seu curso, em semestres, dos alunos que participaram do PBL. Gerado no SPSS.	109
4.5	Distribuição dos alunos nos cursos que participaram do PBL ($n = 61$).	111
4.6	Estrutura da Estatística de um curso de graduação da área da saúde.	112
4.7	Testes de Hipóteses comumente trabalhados em um curso de Bioestatística para a graduação	113
4.8	Entrada do Bioestat dos valores para Exercício 1. Acessível por: Estatísticas > Uma Amostra > Teste Z: Resumo Amostral.	129
4.9	Saída do Bioestat para o Exercício 1.	130
4.10	Saída do Bioestat dos valores para Exercício 2. Acessível por: Estatísticas > Uma Amostra > Teste t: Dados Amostrais.	132

5.1	Cabeçalho e primeiros dados da base de dados do estudo piloto, referido com “Diabete mulheres 21+”.	137
5.2	Escala de Satisfação de Vida.	153
5.3	Dois gráficos de dispersão apresentados pelo grupo A.	156
5.4	Gráfico de dispersão com explicação do grupo B.	157
5.5	Interpretação e gráfico de dispersão do grupo D.	158
5.6	Interpretação do histograma dos escores do grupo B.	159
5.7	Interpretação do histograma dos escores do grupo D.	160
5.8	Interpretação do <i>boxplot</i> dos escores do grupo C.	161
5.9	Interpretação do <i>boxplot</i> dos escores do grupo D.	161
5.10	Definições de desvio padrão (acima grupo B, abaixo grupo D).	161
5.11	Interpretação da variação dos escores da ESV do grupo D.	162
7.1	Medidas resumo para a variável escore. (Gerado por SPSS)	190
7.2	Distribuição de frequências de escore, média, desvio padrão, quantidade de elementos (N) e distribuição normal correspondente. (Gerado no SPSS)	191
7.3	<i>Boxplot</i> para a variável escore com média, desvio padrão, quartis e <i>whiskers</i> . (Gerado no Orange Canvas)	191
7.4	Gráfico do escore versus a proporção de acertos por item. O número 1 refere-se ao item Q1, o 7 ao Q7 e a letra d indica o item Q10. (Gerado pelo pacote 'ltm' do R)	194
7.5	Gráfico do escore versus a proporção de acertos por item. O número 2 refere-se ao item Q2, o 4 ao Q4 e o 9 ao Q9. (Gerado pelo pacote 'ltm' do R)	195
7.6	Médias dos escores separadas por participação no PBL, curso e gênero. (Gerado por Tableau)	195

7.7	<i>Boxplots</i> dos escores com média, desvio padrão, quartis e whiskers, do gênero feminino que participaram (Sim) ou não participaram (Não) do PBL. (Gerado no Orange Canvas)	196
7.8	Curva de estimação da densidade de probabilidade núcleo gaussiano para os escores do gênero feminino, separados por participação no PBL. (Gerado por Orange Canvas)	197
7.9	Proporções de acertos obtidos pelo gênero feminino, em cada questão (Gerado por Tableau).	198
7.10	<i>Boxplots</i> dos escores com média, desvio padrão, quartis e whiskers, do gênero masculino que participaram (Sim) ou não participaram (Não) do PBL. (Gerado no Orange Canvas)	199
7.11	Curva de estimação da densidade de probabilidade núcleo gaussiano para os escores do gênero masculino, separados por participação no PBL. (Gerado por Orange Canvas)	199
7.12	Proporções de acertos obtidos pelo gênero masculino, em cada questão (Gerado por Tableau).	200
7.13	<i>Boxplots</i> dos escores com média, desvio padrão, quartis e whiskers, das diferentes turmas. (Gerado no Orange Canvas)	201
7.14	Curva de estimação da densidade de probabilidade núcleo gaussiano para os escores, separados por turma. (Gerado por Orange Canvas)	202
7.15	Proporções de acertos obtidos pelas diferentes turmas em cada questão (Gerado por Tableau).	203
7.16	Proporções de acertos obtidos pelas turmas do curso de Psicologia separados por participação do PBL (Gerado por Tableau).	204
7.17	<i>Boxplots</i> de escore separado por participação no PBL. (Gerado por Orange Canvas)	205

7.18	Curva de estimação da densidade de probabilidade núcleo gaussiano para os escores separados por participação no PBL. (Gerado por Orange Canvas)	205
7.19	Distribuição dos escores e média dos escores separados por participação no PBL e por Q1, Q2, ..., Q10. (Gerado por Tableau)	206
7.20	Descritivo do escore total estratificado por PBL (Gerado no SPSS)	208
7.21	<i>Boxplots</i> comparativos do escore total estratificados por PBL (Gerado no Excel)	209
7.22	Distribuição de frequências do escore total dentro dos níveis de domínio para quem participou ou não do PBL (Excel)	210
7.23	Descritivo do escore da questão 1 estratificado por PBL (SPSS)	211
7.24	Descritivo do escore da questão 2 estratificado por PBL (SPSS)	212
7.25	Médias das categorias estratificadas por PBL (Excel)	213
7.26	Distribuições dos alunos quanto ao nível atingido nas respostas das questões um e dois juntas, dentre as quatro primeiras categorias, separados pela participação no PBL	215
7.27	Distribuições dos alunos quanto ao nível atingido nas respostas das questões um e dois juntas, dentre as três últimas categorias, separados pela participação no PBL	216
7.28	Uma boa resposta à questão 1, dada pelo estudante ID25 . Os grifos coloridos são anotações do professor/tutor	219
7.29	Uma resposta com problemas conceituais à questão 1, dada pelo estudante ID86. Os grifos coloridos são anotações do professor/tutor	221
7.30	Uma resposta adequada à questão 2 dada pelo estudante ID22 . Os grifos coloridos são anotações do professor/tutor	222
7.31	Uma resposta inadequada à questão 2 dada pelo estudante ID82. Os grifos coloridos são anotações do professor/tutor	223

7.32 Resposta do estudante ID16 à questão 1. Os grifos coloridos são anotações do professor/tutor.	224
7.33 Resposta do estudante ID79 à questão 1.	225
7.34 Resposta do estudante ID87 à questão 2.	225
7.35 Desenvolvimento do estudante ID90 à questão 1. Os grifos coloridos são anotações do professor/tutor.	226
7.36 Desenvolvimento do estudante ID39 à questão 1. Os grifos em vermelho mais claro são anotações do professor/tutor.	227
7.37 Desenvolvimento do estudante ID95 à questão 1. Os grifos em vermelho mais claro são anotações do professor/tutor.	227
7.38 Desenvolvimento do estudante ID81 à questão 2. O grifo em vermelho é do professor/tutor.	228
7.39 Interpretação do estudante ID19 à questão 2.	228
7.40 Resposta do estudante ID76 à questão 1.	229
7.41 Resposta do estudante ID84 à questão 2. Os grifos em vermelho são anotações do professor/tutor.	229
7.42 Resposta do estudante ID43 à questão 2.	230
7.43 Resposta do estudante ID45 à questão 2.	230
7.44 Resposta do estudante ID55 à questão 2. Os grifos em vermelho são do professor/tutor.	231
C.1 Mapa conceitual do campo conceitual de Teste de Hipóteses para a Média Populacional (THMP).	269
C.2 Mapa conceitual do campo conceitual de Teste de Hipóteses para a Média Populacional (THMP). (Cont.)	270
C.3 Mapa conceitual do campo conceitual de Teste de Hipóteses para a Média Populacional (THMP). (Cont.)	271

D.1	Captura de tela do Ambiente Moodle para atividades do estudo piloto.	273
D.2	Materiais de apoio oferecido aos estudantes durante o estudo piloto. . . .	274
D.3	Cabeçalho e primeiros dados da base de dados do estudo piloto.	275
E.1	Escala de Satisfação de Vida.	279
H.1	Exemplo de marcação de respostas na Escala de Autoeficácia	290
H.2	Tabela de conversão de itens negativos na Escala de Autoeficácia	290
H.3	Tabela de normas para interpretação de resultados na Escala de Autoeficácia	291

Lista de Tabelas

2.1	Distribuição entre países de origem e área.	22
2.2	Distribuição das categorias pelas áreas.	22
3.1	Tipos de erro associados à realização dos testes estatísticos e suas respectivas probabilidades	71
3.2	Campo Conceitual dos Teste de Hipóteses: método da região crítica . . .	90
3.3	Campo Conceitual dos Teste de Hipóteses: método do p-valor	91
4.1	Pontuação média em Matemática, de cinco pontos possíveis, na prova do vestibular de verão, dos estudantes ingressantes em 2018 e 2019, nos três cursos. Em 2019, a entrada para os alunos dos cursos de Biomedicina e Nutrição ocorreu por meio de redação. Fonte: Coordenadoria de Seleção e Ingresso da IES.	107
4.2	Distribuição dos TFGs quanto ao grau de uso de análise quantitativa, nos três cursos.	110
4.3	Distribuição dos TFGs quanto ao uso ou não de análise quantitativa, nos três cursos.	110
4.4	Quantidades de alunos por curso, turma e participação no PBL.	119
4.5	Tabela contendo a categoria de análise, o escore atribuído e a descrição de cada nível.	125

4.6	Tabela contendo a categoria de análise, o escore atribuído e a descrição de cada nível (cont.).	126
4.7	Tabela de classificação do nível de domínio do CC dos THMP de acordo com o escore total no instrumento para questões abertas.	127
6.1	Comparação entre o que se considera como metodologia tradicional e a metodologia do PBL apoiada na TCC abordada nesta pesquisa.	174
7.1	Resumo dos resultados auferidos pelo Diário de Classe.	188
7.2	Resumo dos resultados auferidos pelo Diário de Classe (continuação).	189
7.3	Medidas da TCT por item. (Gerado por Excel e R, pacote ltm)	192
7.4	Tabela de classificação do nível de domínio do CC dos THMP de acordo com o escore total das sete categorias do CC dos THMP em duas questões abertas.	209
7.5	Tabela de equivalência para os escores entre zero e dois (tabela 4.5) e zero e quatro (figura 4.5), para cada nível.	214
7.6	Comparação dos resultados quantitativos e qualitativos para o instrumento para determinação do nível de domínio do CC dos THMP, por categoria.	233
7.7	Continuação da tabela 7.6.	234

Sumário

1	Introdução	1
1.1	Motivações da Pesquisa	1
1.2	A Crise Científica dos Testes de Hipóteses	1
1.3	Objetivos de Pesquisa	6
1.3.1	Objetivo Geral	6
1.3.2	Objetivos Específicos	6
2	Revisão Sistemática de Literatura	9
2.1	Introdução	9
2.2	O Ensino de Estatística até 1995: a revisão de Garfield (1995)	11
2.3	Os Caminhos do Ensino de Estatística entre 1995 e 2007	14
2.4	A Revisão Própria de Literatura: o ensino de Estatística de 2007 a 2017	18
2.5	Análise e Categorização	21
2.6	Análise das Categorias	22
2.6.1	Aspectos Psicológicos	22
2.6.2	Metodologias de ensino	31
2.7	Considerações Finais Acerca da Revisão de Literatura	37
3	Fundamentação Teórica	41
3.1	A Teoria dos Campos Conceituais	44

3.1.1	A TCC como guia para a organização do ensino	53
3.2	Raciocínio, Letramento e Pensamento Estatístico	57
3.3	O Conhecimento Científico dos Testes de Hipóteses	64
3.3.1	Estatística, Inferência Estatística e Testes de Hipóteses: uma Abordagem Típica da Área da Saúde	65
3.3.2	Testes de Hipóteses para a Média Populacional	71
3.4	O Campo Conceitual dos Testes de Hipóteses na Área da Saúde	78
3.4.1	O Campo Conceitual dos Testes de hipóteses para a Média Po- pulacional	82
3.5	<i>Problem Based Learning</i> - PBL	94
4	Metodologia de Pesquisa	101
4.1	Contexto da Pesquisa e Caracterização da Amostra	106
4.2	Materiais e Métodos	116
4.2.1	Diário de Classe	116
4.2.2	Instrumento para Análise do Rendimento	117
4.2.3	Instrumento para Análise do Nível de Domínio	123
5	Atividades Prévias	135
5.1	Estudo Piloto	135
5.1.1	Direcionamentos Provenientes do Estudo Piloto	141
5.2	Aplicação Prévia do PBL	145
6	Aplicação das Atividades: A Metodologia de Ensino	163
7	Resultados	177
7.1	Diário de Classe	177
7.1.1	Descrição do Diário de Classe	179
7.1.2	Resultados do Diário de Classe	187

7.2	Análise do Rendimento	190
7.3	Nível de Domínio do CC dos THMP	207
7.3.1	Resultados Quantitativos	207
7.3.2	Resultados Qualitativos	218
7.3.3	Resumo dos Resultados	231
8	Conclusão e Possíveis Implicações	235
8.1	Prolongamentos da Pesquisa	248
A	Plano de Ensino da Bioestatística	261
B	Caracterização dos Cursos de Biomedicina, Nutrição e Psicologia	264
C	Mapa Conceitual do CC dos THMP	268
D	Capturas de Telas do Ambiente Moodle	272
E	A Escala de Satisfação de Vida (ESV)	276
F	Situação e Material de Apoio para Aplicação Definitiva: Curso de Psicologia	283
G	Situação e Material de Apoio para Aplicação Definitiva: Curso de Nutrição	286
H	Outras Situações Envolvendo os THMP	289
I	Instrumento para Análise do Rendimento	293
J	Termo de Consentimento	298

Capítulo 1

Introdução

1.1 Motivações da Pesquisa

A ideia da realização deste trabalho remonta ao ano de 2009, quando comecei a dar aulas de estatística na Universidade Franciscana (à época, chamada Centro Universitário Franciscano), onde fui naturalmente me direcionando para a área da saúde e percebendo uma disparidade entre os conteúdos ensinados e o que se exigia dos futuros profissionais da área da saúde no domínio de estatística.

Em 2016, ingressei no Doutorado em Ensino de Ciências da Universidad de Burgos, e entrei em contato com a Teoria dos Campos Conceituais (TCC) e o *Problem Based Learning* (PBL). Surgiu, então, o seguinte tema de pesquisa: propor e analisar uma experiência de ensino de estatística para a área da saúde, aplicada por meio da PBL e apoiada na TCC.

1.2 A Crise Científica dos Testes de Hipóteses

Em 2014, a revista *Nature* trouxe um artigo de Regina Nuzzo sobre erros estatísticos nas publicações científicas, particularmente dos que se referem à interpretação do p-

valor. Nuzzo conta a experiência de Matt Motyl, aluno de doutorado em Psicologia o qual havia encontrado, após uma pesquisa com 2000 pessoas, que extremistas políticos não percebiam tons de cinza tanto quanto pessoas politicamente moderadas. Motyl informou um p-valor bastante significativo, 0,01 e replicou seu experimento com outras amostras, encontrando um p-valor de 0,51, e o efeito havia desaparecido. O problema não estava nos dados, nem na análise, mas na objetividade e confiabilidade demasiadamente atribuídas ao p-valor.

Conforme Nuzzo (2014, p. 150), quando o estatístico Ronald Aylmer Fisher criou o p-valor, não procurava por uma decisão definitiva, mas uma parte de um todo bem mais amplo de conhecimentos prévios mesclados com dados num processo fluído de pesquisa.

Nuzzo (2014) mostra que a abordagem alternativa ao p-valor, criada pelo matemático Jerzy Neyman e pelo estatístico Egon Pearson, contrastou com a de Fisher, instaurando uma discussão a qual avançou para além do campo das ideias interferindo no trabalho de outros cientistas, em geral não estatísticos, os quais propuseram um método híbrido de análise, que não se aprofunda em nenhuma das abordagens. Um exemplo de uso comum desta abordagem híbrida pode ser visto na figura 1.1.

Nuzzo (2014) destaca que tais problemas teóricos desencadearam atitudes científicas questionáveis, como o *p-hacking*, em que se tentam vários testes estatísticos até encontrarem o p-valor que se deseja, desconsiderando os pré-requisitos dos testes utilizados.

Em última análise, o estudo adequado da estatística em geral e dos testes de hipóteses, em particular, reforçam a ideia de que o pensamento do pesquisador precisa ater-se às questões subjetivas do estudo, confrontando os resultados dos testes com a teoria conhecida até então. Neste sentido, não há resposta somente numérica, quantitativa às questões da área da saúde, senão uma confluência de abordagens.

Nuzzo (2014) salienta que uma mudança de paradigma é necessária e passa pela

	Valor de Teste = 0,045					
	t	df	Sig. (2 extremidades)	Diferença média	95% Intervalo de Confiança da Diferença	
					Inferior	Superior
Livre	5,659	18	,000	,049000	,03081	,06719
Pastosa	2,731	18	,014	,048947	,01130	,08660
Líquida/Pastosa	,142	17	,889	,002500	-,03464	,03964
Líquida	,076	15	,941	,001500	-,04079	,04379

Figura 1.1: Exemplo de saída dada pelo *software* SPSS, com mais de uma abordagem.

forma que a estatística é ensinada, como a análise de dados é feita e os resultados são reportados e interpretados.

Em março de 2019, a revista *Nature* volta a publicar um artigo ressaltando problemas com os testes de hipóteses. Intitulado “*Scientists rise up against statistical significance*”, o texto termina de forma contundente.

Nosso chamado para retirar a significância estatística e usar intervalos de confiança como intervalos de compatibilidade não é uma panaceia. Embora elimine muitas práticas ruins, poderia muito bem introduzir novas. Assim, monitorar a literatura quanto a abusos estatísticos deve ser uma prioridade permanente para a comunidade científica. Porém, a erradicação da categorização ajudará a interromper reivindicações superconfiantes, declarações injustificadas de “sem diferença” e declarações absurdas sobre “falha de replicação” quando os resultados dos estudos originais e de replicação forem altamente compatíveis. O uso indevido da significância estatística causou muitos danos à comunidade científica e àqueles que dependem de aconselhamento científico. P-valores, intervalos e outras medidas estatísticas têm seu lugar, mas é hora da significância estatística desaparecer (Amrhein et al., 2019, p. 307).

O novo artigo da Nature expõe problemas referentes à aplicação e interpretação da estatística. Além desse, trabalhos como os de [Cohen \(2011\)](#) e [Panagiotakos \(2008\)](#) ressaltam problemas com a má interpretação do p-valor. O problema ganhou tamanha proporção que a *American Statistics Assosiation* (ASA), preocupada com o mau uso do p-valor, lançou o “*The ASA’s Statement on p-Values: Context, Process, and Purpose*” ([Wasserstein e Lazar, 2016](#)), para tentar regular o seu uso.

Além disso, [Jones et al. \(2001\)](#); ? apud [Kuwaiti \(2015\)](#), lembram que o treinamento médico, em todo o mundo, está se tornando mais centrado no aluno, com ênfase na aprendizagem ativa e não na obtenção passiva de conhecimento, e [Smith \(2010\)](#) destaca a importância de se diminuir a distância entre o que é necessário e o que se ensina de Bioestatística. Já [Herman et al. \(2007\)](#) salientam que o currículo atual não ensina o que os médicos necessitam, enquanto [Herman et al. \(2007\)](#) enfatizam a importância da promoção da aprendizagem autônoma.

Nesse cenário, a formação básica em estatística precisa de atenção desde o primeiro contato com a disciplina, o que ocorre, normalmente, nos primeiros semestres da graduação. Neste período da formação, não é raro que o sentimento do aluno frente a estatística seja negativo, principalmente sendo ocasionado pela relação com a matemática e/ou do sentimento de que a estatística é descontextualizada da área da saúde, entre outras razões, discutidas em [Fioravanti et al. \(2019\)](#).

Estes autores destacam que o ensino de Estatística, na área da saúde, evoluiu como área de estudo e pesquisas, e os esforços pontuais e particulares passaram a contar com instruções e direcionamentos institucionais específicos. Entretanto, faltam instrumentos para avaliar o que os alunos aprendem, e, além disso, o uso das teorias de aprendizagem é ínfimo, o que impacta em metodologias embasadas somente na prática dos professores, de acordo com [Fioravanti et al. \(2019, p. 68\)](#).

Particularmente no que se refere aos testes estatísticos, [Post e Van Duijn \(2014\)](#) afirmam que é preciso trabalhar para o entendimento dos princípios por trás dos testes

de hipóteses, assim, mudanças no ensino destes precisam ser feitas.

Segundo [Wasserstein e Lazar \(2016\)](#), técnicas escolhidas, análises conduzidas apropriadamente bem como interpretação correta dos resultados estatísticos também desempenham um papel fundamental para garantir que as conclusões sejam sólidas e que a incerteza em torno delas seja representada adequadamente.

Além disso, [Sebastiani \(2011\)](#) refere-se como um obstáculo epistemológico à abordagem dos testes de hipóteses e salienta que "esse assunto ainda carece de trabalhos que tentem mapear quais são as dificuldades específicas enfrentadas pelos Alunos no processo de ensino e aprendizagem".

Neste sentido, metodologias de ensino que privilegiam o aprofundamento do entendimento conceitual dos Testes de Hipóteses para a Média Populacional (THMP) podem auxiliar no bom uso do p-valor e dos demais conceitos subjacentes à testagem de hipóteses. Entretanto, não há consenso quanto a superioridade das metodologias ativas sobre as tradicionais, e parece útil que propostas didáticas ativas sejam comparadas com estas, ao menos quanto ao rendimento dos alunos, por meio de instrumentos que já estão estabelecidos, como as provas de múltipla escolha, por exemplo.

Considerando os relatos supracitados e que a nossa parte para aprimorar este quadro passa pela proposição de melhorias para as classes de estatística das universidades de ciências da saúde, este trabalho traz uma proposta de ensino do conteúdo de Teste de Hipóteses, com foco para a média populacional, por meio da metodologia de ensino chamada *Problem Based Learning* (PBL) e apoiada na teoria dos campos conceituais (TCC) de Gérard Vergnaud.

Isto posto, passa-se a estabelecer os objetivos de pesquisa.

1.3 Objetivos de Pesquisa

Dado o exposto sobre o ensino de estatística para a área da saúde e sobre a Teoria dos Campos Conceituais de Verganud, é possível formular as seguintes questões de pesquisa:

- Em que medida se estabelece o campo conceitual dos testes de hipóteses para a média populacional de alunos da área da saúde, quando expostos ao PBL apoiada pela TCC?
- Relativamente ao conteúdo dos testes de hipóteses para a média populacional, existe diferença significativa entre o rendimento médio de alunos que experienciaram o PBL apoiado pela TCC e os que não o experimentaram?

A partir dessas questões, estabelece-se a seguinte hipótese de pesquisa: *A aplicação do PBL apoiado pela TCC cria um ambiente investigativo acerca de situações da área da saúde, nas aulas de bioestatística, provocando a desconforto dos invariantes operatórios dos alunos e sua readaptação no decorrer do processo, proporcionando um domínio amplo do campo conceitual dos testes de hipóteses para média de uma amostra, a tal ponto que o rendimento obtido pelos alunos seja compatível ou maior do que daqueles que não passaram pela dinâmica do PBL apoiado pela TCC.*

1.3.1 Objetivo Geral

Avaliar e caracterizar o nível de domínio do campo conceitual dos testes de hipóteses para a média populacional de alunos de graduação na área da saúde, quando se utiliza a metodologia PBL apoiada pela TCC.

1.3.2 Objetivos Específicos

- Realizar uma revisão de literatura capaz de colocar em perspectiva o ensino de Estatística, para a área da saúde, produzido até então;

- Propor e analisar uma sequência didática para o ensino de teste de hipóteses para a média populacional, de acordo com o PBL e amparado na TCC;
- Implementar um instrumento de diagnóstico do nível de domínio do campo conceitual do conteúdo de teste de hipóteses para a média populacional, alcançado pelos alunos após a disciplina de estatística para a área da saúde;
- Implementar um instrumento objetivo para a determinação do rendimento dos alunos acerca dos testes de hipóteses para a média populacional;
- Verificar se existe diferença significativa entre a média do rendimento de Alunos que experienciaram o PBL assistido pela TCC e os que não o experienciaram.

Para atingir estes objetivos, esta tese está organizada de acordo com os seguintes capítulos:

- Revisão de Literatura, que promove um resgate do que foi feito acerca do ensino de Estatística, no ensino superior em geral e voltado para a área da saúde, em particular [*Capítulo 2*];
- Fundamentação Teórica balizadora da Teoria dos Campos Conceituais (TCC), do raciocínio, letramento e pensamento estatístico até a delimitação dos Campo Conceitual dos Testes de Hipóteses para a Média Populacional, e culmina na metodologia de ensino *Problem Based Learning* (PBL) [*Capítulo 3*];
- Metodologia de Pesquisa, que ambienta o estudo e estabelece os parâmetros qualitativos e quantitativos capazes de extrair as respostas para as questões de pesquisa [*Capítulo 4*];
- Atividades Prévias, que testaram e permitiram o aprimoramento da primeira versão da metodologia de ensino e pesquisa [*Capítulo 5*];

- Metodologia de Ensino, que, fundamentada no PBL e apoiada na TCC, compôs o ambiente de ensino que levantou os dados desta pesquisa [*Capítulo 6*];
- Resultados, onde se analisam os dados obtidos a partir da metodologia apresentada no capítulo anterior [*Capítulo 7*];
- Conclusões e Prolongamentos da Pesquisa apresenta os excertos que a experiência permitiu inferir, além dos próximos passos que se seguirão como decorrência desta tese [*Capítulo 8*];
- Anexos, não é um capítulo, mas abarca os materiais utilizados na pesquisa.

Capítulo 2

Revisão Sistemática de Literatura

2.1 Introdução

Esta revisão bibliográfica agrupa questões relativas ao ensino de Estatística para a área da saúde, levantadas desde o final da década de 70 até o ano de 2017 e seus achados podem também serem visitados em [Fioravanti et al. \(2019\)](#).

Partindo das revisões de literatura sobre o ensino de Estatística em geral, realizadas por Garfield (1995) e Garfield e Ben Zvi (2007), em seguida, será apresentada a revisão própria de literatura sobre o ensino de Estatística para a área da saúde, compreendida no período de 2007 a 2017. As revisões de Garfield expõem uma ideia geral dos caminhos percorridos pelo ensino da Estatística, principalmente no Ensino Superior, permitindo que se infira um contexto sobre o tema, no qual se insere o ensino de Estatística para a área da saúde, foco da revisão própria de literatura.

O objetivo desta revisão é delinear um panorama sobre o que foi investigado acerca do ensino de Estatística para a área da saúde (tanto em termos de dificuldades quanto em termos das metodologias didáticas propostas para melhorar sua aprendizagem), a

partir do que fora publicado em artigos científicos, tanto da área da saúde quanto do ensino de Matemática e de Estatística, formando parte de uma pesquisa em andamento que visa aprimorar o ensino de Estatística para a área da saúde. Os resultados desta revisão podem auxiliar os pesquisadores sobre o que falta investigar e quais linhas se pode dar continuidade.

A relevância do tema se deve à necessidade de qualificação do ensino de Estatística, quando se consideram os contextos atuais da medicina baseada em evidências e da epidemiologia, por exemplo. A prática baseada em evidências (Evidence-Based Practice - EBP), segundo McKibbin (1998), é uma abordagem para os cuidados de saúde, na qual os profissionais utilizam a melhor evidência disponível para tomar decisões clínicas para pacientes individuais. Os cuidados de saúde são individualizados e estão em constante mudança, envolvendo incertezas e probabilidades, aspectos inerentes ao raciocínio estatístico. Por outro lado, conforme Bonita, Beaglehole e Kjellstrom (2010), os conceitos e ferramentas de Estatística são necessários para a síntese e análise dos dados e os estudos epidemiológicos requerem o uso de amostras para que sejam feitas inferências sobre uma população. Além disso, Tishkovskaya e Lancaster (2012) ponderam que a Estatística tem sido e continuará a ser um dos temas mais ensinados no nível universitário.

Weissgerber et al. (2016) salientam que efeitos negativos da ineficácia do ensino de Estatística já são sentidos no meio acadêmico, visto que artigos com problemas metodológicos, conhecimento estatístico insuficiente e impossível reprodutibilidade, têm sido encontrados.

A Estatística faz parte da formação do profissional da saúde e os estudantes podem estar se posicionando de forma negativa perante esta demanda.

O letramento estatístico é uma habilidade necessária para que os médicos compreendam e implementem pesquisas. Embora um bioestatístico deva ser consultado antes e durante todo o curso da maioria dos esforços da pesquisa,

é imperativo que o clínico que conduz a pesquisa compreenda os princípios da análise Estatística. (Ing, 2016, p. 142)

Dado que o ensino da Estatística no Ensino Superior para a área da saúde encontra-se inserido no ensino da Estatística no Ensino Superior em geral, opta-se por começar este trabalho com os resultados obtidos em duas abrangentes revisões dessa área. Inicia-se com a análise do artigo de Garfield (1995), que mapeia os caminhos do ensino de Estatística até então.

2.2 O Ensino de Estatística até 1995: a revisão de Garfield (1995)

Em 1995, Joan Garfield publicou na revista *International Statistical Review*, uma revisão de literatura intitulada “Como os estudantes aprendem Estatística”, sobre o ensino de Estatística em nível universitário, com o argumento de que os educadores precisam conhecer o que realmente esperam que os estudantes aprendam, a fim de modificar a forma como ensinam, utilizando a avaliação para determinar se estão sendo efetivos e se os estudantes estão desenvolvendo a compreensão e a competência Estatística. Segundo o Google Scholar, seu artigo possuía 607 citações, até quatorze de janeiro de 2018.

Garfield (1995) destaca a importância do conhecimento das teorias de aprendizagem e descreve que o construtivismo, derivado das teorias de Piaget, concebe a aprendizagem com uma construção ativa, por parte do estudante, do seu próprio conhecimento. No cenário construtivista, o conhecimento não é transmitido ou repassado, mas construído a partir de um ambiente adequado, proporcionado pelo professor.

A autora questiona se o que se tinha feito nas aulas proporcionaria que os alunos atingissem tais objetivos, como pode ser verificado na citação a seguir:

Muitas classes de Estatística da faculdade consistem em assistir às explicações e fazer as tarefas em livros didáticos ou em laboratórios de informática.

Essas atividades ajudam a atingir os objetivos para nossos alunos? Os alunos estão adequadamente preparados para utilizarem o pensamento e o raciocínio estatísticos, coletar e analisar dados, escrever e comunicar os resultados da resolução de problemas estatísticos reais? Muitas pesquisas indicam que os estudantes não estão aprendendo o que queremos. (Garfield, 1995, p. 27)

Garfield (1995) classifica seus achados em três categorias: pesquisa psicológica, pesquisa em educação estatística e pesquisa em educação matemática (seções quatro, cinco e seis, respectivamente).

Em relação à pesquisa psicológica, a autora destaca que a maior parte das pesquisas atenta para como adultos compreendem, ou não compreendem, ideias Estatísticas particulares. Esta heurística revela alguns pensamentos prevalentes e inconsistentes, como a falácia do jogador, em que muitas pessoas acreditam que depois de uma grande sequência de caras é mais provável que a próxima jogada seja uma coroa. A autora identificou problemas de entendimento relativos à correlação e à causalidade, à probabilidade condicional, à independência, à aleatoriedade e à média ponderada. Garfield (1995) salienta que raciocínios estatísticos inapropriados são generalizados e persistentes, ocorrem em todos os níveis de idade, inclusive entre pesquisadores experientes.

A pesquisa relativa à educação estatística estava menos focada em padrões de pensamento e mais em como a Estatística é aprendida. Neste sentido, a autora compilou uma série de atributos que ajudam os estudantes a aprender:

- Cursos baseados em atividades (*Activity-based courses*) e o uso de pequenos grupos parecem ajudar os estudantes a superar alguns equívocos em probabilidade e aumentar o aprendizado de conceitos estatísticos;
- Quando os alunos são avaliados e recebem os resultados desta avaliação, seguidos de atividades de correção, nas quais são encorajados a explicar e conjecturar soluções e verificar se o que fizeram faz sentido, estas atividades ajudam a superar equívocos;

- Ideias sobre probabilidade de amostras melhoram quando é preciso fazer previsões antes da tomada dos dados e comparações com o resultado experimental obtido;
- Simulações computacionais melhoram as respostas dos alunos sobre variabilidade e problemas sobre probabilidade;
- Utilizar softwares que permitem a visualização e interação com os dados, melhora o entendimento dos estudantes sobre a aleatoriedade e análise de dados.

A autora destaca os achados provenientes da educação Matemática:

- Investir mais tempo no entendimento de como um algoritmo funciona, quais as habilidades envolvidas e como um conceito difere de outro, melhoram a capacidade do estudante em resolver problemas;
- Trabalhar com grupos pequenos melhora a produtividade de toda a turma, suas atitudes e suas conquistas;
- Trabalhar com exercícios resolvidos pode ser mais efetivo do que propor listas de exercícios convencionais, como os temas de casa;
- Os estudantes aprendem mais com problemas abertos do que com problemas de uma única resposta;
- Atividades matemáticas do tipo “Escrevendo para aprender” (*“Writing to learn”*) parecem auxiliar no entendimento matemático;
- Metodologias inovadoras como resolução de problemas e pensamento superior (*“higher order thinking”*) parecem melhorar o rendimento dos estudantes, sem sofrer qualquer perda em testes tradicionais.

Garfield (1995) propõe dez princípios de aprendizagem Estatística, construídos de acordo com seus achados e no contexto do construtivismo. Estes princípios foram recompilados em oito e são apresentados no próximo item desta revisão.

Em resumo, Garfield (1995) destaca que o ensino de Estatística pode ser mais efetivo se os professores determinarem o que eles realmente querem que os estudantes saibam e façam ao final do seu curso, providenciando atividades desenhadas para este fim. Além disso, a autora considera a importância da realização de avaliações incorporadas ao processo de aprendizagem para que alunos e professores determinem se os objetivos foram alcançados, isto em tempo de rever alguma estratégia antes do final do curso.

Os docentes precisam avaliar o que foi encontrado nas pesquisas e determinar como irão adaptar com a sua realidade particular, precisam analisar e redefinir continuamente suas teorias de como os estudantes aprendem Estatística. Os discentes devem ser encorajados a avaliar seu próprio aprendizado e as noções de como aprendem pensando sobre o processo de ensino-aprendizagem.

Diante disso, Garfield (1995) conclui ponderando que a maioria das pesquisas consultadas trazia somente implicações gerais e que muito havia a ser aprendido sobre problemas particulares acerca do uso do computador, das técnicas para superar concepções errôneas, do tipo de atividades em pequenos grupos e de quais materiais avaliativos melhor informam o professor sobre o entendimento dos alunos.

A revisão de Garfield (1995) trouxe uma ampla visão do estado do ensino de Estatística para alunos de graduação, identificou linhas de pesquisa, propôs melhorias e indicou questões a serem respondidas. Dez anos depois, a autora volta a analisar o estado da arte do ensino de Estatística, no trabalho descrito a seguir.

2.3 Os Caminhos do Ensino de Estatística entre 1995 e 2007

Garfield e Ben-Zvi (2007) revisitam o trabalho de Garfield (1995), a fim de determinarem os caminhos tomados pelo ensino de Estatística até 2007. Esta nova revisão possuía 341 citações em janeiro de 2018, segundo o Google Scholar.

Os autores citam 151 trabalhos, mostrando a expansão do ensino de Estatística quando comparado aos 40 artigos da primeira revisão, entretanto, este artigo aborda o ensino de Estatística em todos os níveis, enquanto o artigo de Garfield (1995) focava no ensino universitário. Além disso, Garfield e Ben-Zvi (2007) resumizam estudos conduzidos por pesquisadores de outras áreas.

Aqui, as ideias principais do artigo de Garfield e Ben-Zvi (2007) são mostradas dando destaque ao Ensino Superior. Os autores iniciam destacando o crescimento das conferências e publicações dedicadas à pesquisa em educação estatística mesmo que, em 2007, ainda se considerava a área como nova, emergente, difícil de localizar suas bases e evoluir a partir delas, pois seus estudos estavam diluídos em publicações de diversas outras disciplinas.

Relativamente às conferências e publicações especializadas, em 2002 surge a *Statistics Education Research Journal*, primeira revista dedicada a publicações de alta qualidade no ensino de Estatística. Em seguida, surgiram: *The International Conference on the Teaching Statistics*, o *International Group for the Psychology of Mathematics Education*, *The Mathematics Education Research Group of Australasia*, *The International Congress on Mathematics Education meetings* e o *International Statistical Institute*, que também relatam pesquisas na área de educação Estatística. A grande quantidade de apresentações e publicações, a partir destas conferências, revela a existência de um grupo ativo de educadores, psicólogos e estatísticos envolvidos no ensino e na aprendizagem da Estatística, além disso, mais de 44 teses doutorais abordaram o tema, segundo Garfield e Ben-Zvi (2007).

Na revisão de Garfield e Ben-Zvi (2007), as pesquisas focaram em métodos de treinamento individual para alcançar um aprimoramento do raciocínio. Conforme os autores, estas pesquisas consideraram que a causa para tantas concepções errôneas encontradas era a inabilidade do uso do raciocínio proporcional, requerido por muitos problemas envolvendo probabilidade, entre outros tipos de erros dentro dos conteúdos

de Tabelas de Contingência, Distribuição Amostral, Testes de Significância e uma variedade de erros de raciocínio estatístico. Destes estudos, os autores depreenderam que as concepções errôneas sobre a Estatística estão disseminadas e são similarmente persistentes em todos os níveis de ensino, até mesmo entre pesquisadores experientes, tal como Garfield (1995) já havia destacado.

Os autores ressaltam que as pesquisas examinaram atividades particulares ou intervenções e alguma ferramenta tecnológica ou método de ensino. A maioria destes estudos envolveu as classes dos próprios pesquisadores, por vezes examinando uma classe, por outras, olhando para várias classes da mesma instituição.

Um grande problema observado em estudos quantitativos foi a falta de instrumentos consistentes e de alta qualidade para avaliar os resultados de aprendizagem dos alunos, sendo comum a utilização das notas (rendimento) como medida, uma vez que estas não trazem evidências de validade e confiabilidade e normalmente não medem resultados generalizáveis.

No entanto, estes estudos revelam que é difícil determinar o impacto de uma determinada ferramenta, método de ensino ou instrução sobre a aprendizagem dos alunos em um curso devido a limitações no desenho do estudo ou avaliação utilizados. Enquanto professores gostariam de contar com estudos de investigação para convencê-los que um método de ensino particular ou ferramenta instrucional leva à melhorar significativamente os resultados dos alunos, esse tipo de evidência não é realmente disponível na literatura de pesquisa. Os resultados de muitos dos estudos comparativos são geralmente limitados a esse tipo de curso particular e não pode ser generalizado para outros cursos. (Garfield & Ben-Zvi, 2007, p. 379)

No contexto das ferramentas tecnológicas, incluindo as de simulação, os autores afirmam que uma ferramenta bem concebida não garante um ensino eficaz. Estudos também mostraram que não ficam evidentes as diferenças de aproveitamento em distintas modalidades de ensino, presencial, on-line ou híbridas.

O que ficou evidente para Garfield e Ben-Zvi (2007) foi que quando os estudantes trabalham de forma colaborativa, em grupos, e considerando suas concepções prévias, há

melhora em seus rendimentos. Alguns estudos encontraram bons resultados quando proporcionaram a evolução do pensamento estatístico informal para o formal, considerando a não linearidade da construção do raciocínio estatístico durante o curso, entretanto, é difícil estabelecer qual método é melhor para o Ensino de Estatística.

Neste sentido, os autores Garfield e Ben-Zvi (2007) pontuam que “desenvolver uma compreensão profunda de conceitos de Estatística é bastante desafiador e não deve ser subestimado”. O processo é demorado, precisa ser bem planejado quanto à trajetória a percorrer, bem como as ferramentas a utilizar, a escolha das questões para discussão e a forma de proporcionar a revisitação destas ideias. Ressaltam, ainda, a importância dos estudos que focam no desenvolvimento de tópicos específicos de Estatística, pois podem auxiliar os professores no entendimento dos erros e dificuldades de compreensão comuns àquele tópico, prevenindo-os.

Finalmente, os autores apresentam seus oito princípios para o aprendizado de Estatística, reestruturados, a partir dos dez princípios propostos por Garfield (1995):

- i Estudantes aprendem pela construção do conhecimento;
- ii Estudantes aprendem pelo envolvimento ativo em atividades de aprendizagem;
- iii Estudantes aprendem a fazer bem somente o que praticam;
- iv É fácil subestimar a dificuldade que os estudantes têm em entender conceitos básicos;
- v É fácil superestimar quão bem os estudantes entendem conceitos básicos;
- vi A aprendizagem melhora se os estudantes tomam consciência e confrontam seus erros de raciocínio;
- vii Ferramentas tecnológicas devem ser utilizadas para ajudar os estudantes a visualizar e explorar dados, não apenas para seguir algoritmos com finais pré-determinados;

viii Estudantes aprendem melhor se recebem *feedbacks* consistentes que auxiliem em sua performance.

Para Garfield e Ben-Zvi (2007) há a necessidade do desenvolvimento de projetos de pesquisa colaborativos, preferencialmente partindo de diferentes disciplinas, a fim de combinar suas experiências e comparar resultados obtidos em muitas instituições, usando instrumentos de medida de alta qualidade.

De maneira geral, o artigo de Garfield (1995) propôs bases para o ensino de Estatística, considerando o que se espera que os alunos aprendam e as suas atitudes com relação à disciplina. Já o artigo de Garfield e Ben-Zvi (2007) destaca o esforço das pesquisas em identificar maneiras mais particulares do ensino e da aprendizagem de Estatística. Neste sentido, observa-se um direcionamento do geral para o particular no foco das pesquisas da área. Contudo, ambos os períodos (até 1995 e de 1995 a 2007) evidenciaram a necessidade de questionamento do ensino tradicional e a apropriada utilização do construtivismo. No entanto, enquanto a revisão de 1995 estava mais dirigida aos próprios professores que ministram estatística, o artigo de 2007 toma certo distanciamento e é mais focado à pesquisa e aos pesquisadores.

Em seguida, apresenta-se o levantamento do estado do ensino de Estatística para a área da saúde, por meio da revisão própria de literatura. Dado que o foco desta revisão está no ensino, usaremos os princípios do aprendizado da Estatística resultantes destas duas revisões como marco.

2.4 A Revisão Própria de Literatura: o ensino de Estatística de 2007 a 2017

Neste ponto, será tratada a revisão própria de pesquisas publicadas sobre o ensino de Estatística para o nível superior e, especificamente, na área da saúde. Para isso, o estudo concentrou-se em artigos completos divulgados exclusivamente, em revistas e eventos

que contam com revisão por pares, contemplando somente trabalhos publicados entre 2007 e 2017.

Foram utilizadas ferramentas de busca on-line contidas em repositórios de artigos das áreas da Estatística (ISI), Ensino de Estatística (IASE) e Ensino de Matemática (RELIME), além de ferramentas de busca científica (Web of Science, Scopus e Scielo). Também foram utilizadas ferramentas automáticas de alerta por e-mail, que monitoram as publicações mais recentes, considerando palavras-chaves pré-definidas da Google Scholar, PubMed e ScienceDirect. Em cada ferramenta de busca, foram utilizadas as seguintes expressões (bem como combinações das palavras contidas em tais expressões): “Statistics Teaching” ou “Teaching Statistics” AND University Level”. Em seguida, a pesquisa sofreu um refinamento de acordo com os filtros que dispunha para restringir os resultados em relação ao tema, como, “Health Care”, “Science Technology” ou “Educational Research” e quanto ao tipo de documento, visto que somente considerou-se artigos em Língua Inglesa, Portuguesa ou Espanhola. Embora existam muitas expressões que remetem ao trabalho estatístico, tais como “Data Analysis” ou “Data Science”, esta revisão não as utilizou, pois estas ainda são pouco usadas na área da saúde de forma relacionada com o ensino de Estatística. O levantamento ocorreu durante janeiro de 2016 a dezembro de 2017. Inicialmente, foram encontrados 343 artigos na Web of Science; 202 na Scopus, e 29 na Scielo. Assim, tendo em conta a repetição de artigos e a busca automática da PubMed, Science Direct e Google Scholar, obtivemos um universo aproximado de 460 artigos.

Os artigos selecionados provêm das seguintes revistas e eventos: Currents in Pharmacy Teaching and Learning, Avaliação e Políticas Públicas em Educação, Global Journal for Research Analysis, Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria, ICOTS 8 (2), Journal of Statistics Education (2), Nurse Education Today, PLOS ONE (2), Preventive Medicine, Probabilidad Condiçionada: Revista de didáctica de la Estadística (2), Psicologia: Ciência e Profissão,



Figura 2.1: Estrutura metodológica da revisão própria de literatura

Psicothema, Revista Brasileira de Educação Médica, Revista Humanidades Médicas, Revista Liberabit, SERJ, Statistics in Medicine, BMC Medical Education, Anales de Psicologia, Clinical Oncology e AHPE.

Os artigos foram pré-selecionados, dentre os filtrados pelas ferramentas, por meio da análise de seus títulos, onde se buscou uma indicação de que se trata de um trabalho sobre ensino de Estatística para a área da saúde no Ensino Superior. Não havendo tal indicação, o artigo era descartado e, se houvesse alguma dúvida sobre seu conteúdo, passava-se a analisar o seu resumo, persistindo a dúvida, lia-se o artigo completo. Ao final do processo, foram selecionados 26 artigos que compuseram a amostra, artigos que foram analisados na íntegra. Apesar da forma sistemática de busca de artigos, esta pesquisa não pode ser definida como uma pesquisa sistemática da literatura. Na 2.1, pode-se verificar a estrutura metodológica da revisão própria de literatura.

2.5 Análise e Categorização

Os trabalhos analisados tratam do ensino de Estatística para a área da saúde, a partir de abordagens não aplicadas e aplicadas. As não aplicadas transitam por uma diversidade de temas que vão da análise das atitudes dos alunos frente à Estatística, passando pelo currículo, avaliação e sugestão de metodologias de ensino, porém, sem a efetiva aplicação de metodologia de ensino aos alunos. Os práticos testam propostas de ensino em turmas da área da saúde. Os artigos encontrados têm origem nas duas grandes áreas as quais se dedica esta revisão: a de ensino (10) e a da saúde (16). Por área de ensino entende-se as revistas e eventos da área de ensino e por área da saúde entende-se as revistas e eventos de próprios da área da saúde que eventualmente tratam do ensino da Estatística. A maior quantidade de trabalhos provenientes da área da saúde em muito se justifica pelo volume de trabalhos que abordam os aspectos psicológicos envolvidos no ensino de Estatística para a área da saúde.

A maior quantidade de trabalhos provenientes da área da saúde, em muito se justifica pelo volume de trabalhos que abordam os aspectos psicológicos envolvidos no ensino de Estatística para a área da saúde.

Relativamente aos países de origem, os artigos apresentam grande espalhamento ao redor do mundo, indicando que as preocupações são semelhantes. Este fato pode ser constatado na [2.1](#).

Os artigos analisados parecem adequar-se em duas categorias, denominadas de aspectos psicológicos e metodologias de ensino, conforme a [2.2](#). A categoria “aspectos psicológicos” contempla os trabalhos que afetam diretamente o ensino de Estatística, mas não através de uma proposta didática, como a ansiedade estatística e as atitudes dos alunos frente à disciplina. Já a categoria “metodologias de ensino” refere-se aos trabalhos dedicados à melhoria da relação pedagógica entre os estudantes, professores e instituições, através de diversas propostas. Dentre as propostas, algumas aplicadas

País	ensino	saúde	Total
Arábia Saudita		1	1
Argentina	1	1	2
Austrália		1	1
Brasil	2	2	4
Caribe	1		1
Espanha	1	2	3
Estados Unidos	3	2	5
Grécia		1	1
Israel		1	1
Paquistão		1	1
Portugal	1		1
Sérvia		2	2
UK	2	1	3
Total	11	15	26

Tabela 2.1: Distribuição entre países de origem e área.

Categorias	Ensino	Saúde	Total
Aspectos psicológicos	4	9	13
Metodologia de ensino	6	7	13
Total	10	16	26

Tabela 2.2: Distribuição das categorias pelas áreas.

e avaliadas e outras não, encontram-se tanto propostas para a sala de aula particular quanto para um currículo de curso.

2.6 Análise das Categorias

2.6.1 Aspectos Psicológicos

Ao que parece, as dificuldades próprias do conteúdo estatístico não constituem o único desafio a ser enfrentado pelos professores de Estatística, existem dificultadores externos que impactam decisivamente no fazer pedagógico dos docentes: a ansiedade estatística e a atitude negativa dos alunos frente à Estatística. Estresse, desmotivação e repulsa dos alunos por não verem ligação com a área da saúde ou porque as aulas não são

estimulantes, como relatam Oliver, Galiana, Cebrià e Sancho (2014) e Espindola, López, Miranda, Ruiz e Díaz (2014). A ansiedade estatística constitui uma emoção caracterizada por extensa preocupação, pensamentos intrusivos, desorganização mental, tensão e ativação fisiológica quando os estudantes estão expostos aos conteúdos, aos problemas, às situações de sala de aula ou contextos avaliativos que envolvam a Estatística (Zeidner, 1990).

Doze artigos abordaram aspectos psicológicos frente à Estatística (ansiedade Estatística e atitudes frente à Estatística) e, após a leitura completa dos trabalhos, foram descritos como pode ser observado na tabela abaixo.

Interpretação própria dos artigos na categoria Aspectos psicológicos.

Autores	Área	Objetivo	O que foi medido	Como foi medido	Metodol. de Pesquisa	Metodol. de Ensino	Teoria de Apend.
(León Bologna & Vaiman, 2013)	Ensino	Verificar se exposição à Estatística em níveis de ensino anteriores melhora a atitude na graduação de Psicologia	Atitude	Escala própria tipo Likert	Quanti+ Multivariada	Não se aplica	Não se aplica
(Turik, Viali, & Moraes, 2012)	Ensino	Identificar os fatores que melhoram a atitude dos alunos frente à estatística	Atitude	Escala de atitudes	Quanti+ TRI+ Multivariada	Não se aplica	Não se aplica
(Stanisavljevic et al., 2014)	Saúde	Medir as atitudes frente à estatística usando uma versão da escala SATS	Atitude	Versão da SATS	Quanti+ Multivariada	Não se aplica	Não se aplica

(Moreira Junior, Zanella, Lopes, & Seidel, 2015)	Saúde	Medir a satisfação dos alunos quanto à Estatística	Atitude	Escala própria tipo Likert	Quanti+ TRI+ Multivariada	Não se aplica	Não se aplica
(Hannigan, Hegarty, & McGrath, 2014)	Saúde	Identificar aspectos que melhoram a atitude dos alunos frente à Estatística	Atitude	ATS, SAS, SATS e SATS-36	Quanti+ Descritiva+ Correlação Multilinear	Não se aplica	Não se aplica
(Vigil-Colet, Lorenzo-Seva, & Condon, 2008)	Saúde	Desenvolver uma escala própria de medida da ansiedade estatística e compara-la com escalas gerais	Ansiedade	Desenvolvimento da SAS	Quanti+ Multivariada	Não se aplica	Não se aplica
(Hernandez, Santos, Silva, Mendes, & Ramos, 2015)	Saúde	Validar a SAS	Ansiedade	SAS	Quanti+ Multivariada	Não se aplica	Não se aplica
(Escalante Gómez, Repetto, & Mattinello, 2012)	Saúde	Avaliar atitudes frente à Estatística	Atitude	SATS	Quanti+ Multivariada	Não se aplica	Não se aplica
(Oliver, Sancho, Galiana, & Cebrià i Sancho, 2014)	Saúde	Validade a SAS	Ansiedade	SAS	Quanti+ Multivariada	Não se aplica	Não se aplica
(Kiekkas et al., 2015)	Saúde	Avaliar as atitudes frente à Estatística de alunos de Enfermagem	Atitude	SAS-36	Quanti+ Inferencial	Não se aplica	Não se aplica
(Willie, Ferguson, Tulloch-Reid, & Mccaw-Binns, 2012)	Ensino	Verificar quais técnicas os alunos preferem, entre centradas no Professor ou mais ativas.	Atitude	Escala própria tipo Likert	Quanti+ Q Análise de Conteúdo+ Quantidades absolutas e relativas	Cita técnicas de ensino diversas sem aplicá-las	Não utilizaram
(Mahboob, Wajid, & Iqbal, 2015)	Saúde	Avaliar a percepção dos alunos frente a uma oficina obrigatória de estatística.	Atitude	Escala própria tipo Likert	Quanti+ Inferencial	Não se aplica	Não se aplica

(Espindola Artola, López Benítez, Miranda Carbonell, Ruiz Socarrás, & Díaz García, 2014)	Saúde	Avaliar os benefícios de metodologias participativas para diminuir o estresse frente a Estatística	Ansiedade	Escala SISCO	Quali+Quantitativa	Metodologia Participativa	Vigotsky
--	-------	--	-----------	--------------	--------------------	---------------------------	----------

Quatro artigos focam a ansiedade estatística, mas alguns dos demais também aportam ideias sobre o tema: Espindola Artola et al. (2014) encontraram uma diminuição na ansiedade Estatística quando se utiliza uma metodologia participativa. Este estudo traz uma proposta didática objetivando a diminuição da ansiedade Estatística e não uma metodologia de ensino focada no aprendizado de Estatística propriamente dito, por isto se classifica na categoria aspectos psicológicos. Hernandez, Santos, Silva, Mendes e Ramos (2015) encontraram um nível de ansiedade mais elevado nas mulheres. Turik, Viali e Moraes (2012) concluíram que estudos estatísticos contextualizados à realidade do estudante contribuem para a diminuição da ansiedade estatística. León e Vaiman (2013) verificaram que a exposição à Estatística na escola melhora o desempenho na faculdade, mas não necessariamente o sentimento que se tem da disciplina. Vigil-Colet, Lorenzo-Seva e Condon (2008) determinaram que os resultados específicos de ansiedade estatística têm uma relação significativa com o desempenho acadêmico nesta disciplina, enquanto medidas mais amplas de ansiedade não os tem.

Estrada, Batanero e Lancaster (2011) citam três instrumentos de medida da ansiedade estatística:

Três dos mais utilizados instrumentos de medida das atitudes em relação à Estatística são Wise (1985) Attitudes Towards Statistics scale (ATS), Roberts e Saxe (1982) Statistics Attitude Survey (SAS) e Chau, Stevens, Dauphine e del Vecchio (1995) Survey of Attitudes Toward Statistics (SATS). (Estrada, Batanero & Lancaster, 2011 p. 164)

Outros instrumentos surgiram, alguns adaptados dos citados acima, como o SATS-

36 de Nowacki (2015) e outros utilizando a Teoria da Resposta ao Item (TRI) em sua análise, como em Moreira Junior et al. (2015).

Os trabalhos desta categoria trouxeram uma forte análise Estatística multivariada tanto para ansiedade quanto para a atitude, normalmente utilizando o alfa de Chronbach e a análise fatorial aplicados sobre questionários do tipo Likert, em geral com 5 níveis e análise estatística por meio da Teoria da Resposta ao Item (TRI). Costumam acessar grande número de amostras, dada a facilidade de aplicação dos questionários, inclusive com versões online. Mesmo que as medidas tenham mais de 20 anos, muitos estudos ainda buscam indícios de validade e não há uma escala que suplante as demais, evidenciando a dificuldade que este tipo de instrumento encontra para afirmar-se. Cabe ressaltar que a maioria (10) tem sua origem na área da saúde, enquanto que os demais estudos (3) se originam da área do ensino.

Um destacado instrumento para a ansiedade é a SAS (*Statistical Anxiety Scale*), criada por Vigil-Colet et al. (2008), que procura estabelecer uma escala estritamente relacionada à ansiedade estatística e pequena o bastante para ser facilmente administrada individualmente ou em uma classe de alunos. O processo de validação da SAS contou com 159 graduandos espanhóis do curso de Psicologia, sendo 139 mulheres e 20 homens, com uma média de 21,6 anos de idade e um desvio padrão de 3,5 anos. Os alunos responderam voluntariamente a três questionários, o primeiro foi a SAS e os demais eram relativos a traços de personalidades. Foi utilizada a análise exploratória fatorial, extraíndo-se três dimensões, usando o método de análise de fator de classificação mínima e o método de rotação mínima Promim. A análise fatorial apresentou três dimensões correlacionadas entre si e que poderiam ser consideradas como subescalas relacionadas de uma escala global. Além disso, a análise paralela foi calculada para analisar a matriz de correlação inter-fator e sugeriu que havia uma dimensão subjacente a ela. Se as três subescalas fossem ansiedade de exame, ansiedade ao pedir ajuda e ansiedade ao interpretar, a escala global seria a ansiedade estatística, conforme Vigil-Colet

et al. (2008) reforçado por Oliver et al. (2014) e Hernandez et al. (2015).

A SAS foi aplicada a alunos italianos e espanhóis (Vigil-Colet et al., 2008), (Oliver et al., 2014) e brasileiros (Hernandez et al., 2015), e se percebe uma incidência maior de ansiedade estatística nas mulheres.

Em relação às atitudes, perante o estudo da Estatística, foram encontrados nove trabalhos. Hannigan, Hegarty e McGrath (2014) apresentam uma escala para medir as atitudes dos estudantes frente ao estudo da estatística. Esta escala foi aplicada em alunos entrantes na faculdade de Medicina, visto que estes estudantes provavelmente trazem experiências prévias do raciocínio quantitativo, oriundas da Matemática e não da Estatística. Os autores aplicaram o instrumento SATS-36 a 121 alunos de primeiro ano logo nas primeiras semanas do semestre, para garantir que nunca tenham sido expostos à Estatística, capturando assim, suas atitudes em relação à disciplina no início da faculdade.

Na amostra, 66% dos respondentes tinha menos de 25 anos, 58% eram mulheres, 59% eram irlandeses e 31% norte americanos, 61% teve uma educação básica em ciências ou engenharia, 85% teve um curso quantitativo na educação básica, com mediana de dois cursos.

A escala SATS-36 foi escolhida por ter sido usada anteriormente, em estudo sobre as atitudes em relação à Estatística de alunos de pós-graduação, permitindo uma análise comparativa entre as pesquisas. A escala mediu seis componentes referentes à atitude: Sentimental, Cognitivo, Utilidade (Valor), Dificuldade, Interesse e Esforço.

A SATS-36 conta com uma versão pré e outra pós-instrução, todavia, os autores utilizaram somente a versão pré-instrução. As 36 sentenças possuem uma escala de sete itens, na qual o primeiro refere-se a “discordo fortemente” e o sétimo a “concordo fortemente”. Dentre as 36 sentenças, 19 estão valoradas de maneira inversa e precisaram ser recodificadas, desta forma, quanto mais próximo de 7 as médias das respostas estão, mais positiva é a atitude em relação à Estatística.

Hannigan et al. (2014) introduziram, na análise das atitudes, questões relativas à performance dos alunos em Matemática. Utilizaram uma escala de 1 (um) a 7 (sete), na qual 1 (um) representa uma performance muito fraca, enquanto 7 (sete) representa uma performance muito forte. Também perguntaram sobre a quantidade de módulos de Matemática ou Estatística que tiveram em sua educação básica e qual seria a chance de se matricularem em algum curso de Estatística, se tivessem a oportunidade (respostas de 1 (um) a 7 (sete), em que 1 (um) representa nenhuma chance e 7 (sete), muita chance).

A análise dos resultados contou com o cálculo do alfa de Cronbach para medir a consistência interna das respostas dos 36 itens do questionário SATS (alfa = 0,93) e dos seus seis componentes: Valor e Esforço (alfa = 0,79), Dificuldade (alfa = 0,81), Afetivo (alfa = 0,85) e Cognitivo e Interesse (alfa = 0,88). Os valores de alfa obtidos indicam boa confiabilidade dos componentes e são similares aos reportados em outros estudos. A regressão linear multivariada foi usada para prever os escores utilizando variáveis demográficas: idade (< 25 , ≥ 25 sexo, nacionalidade (Irlandês ou não Irlandês) e variáveis representando as experiências educacionais anteriores: número de módulos de estudo quantitativo na educação básica (nenhum, um ou mais) e percepção da sua performance prévia em Matemática (escala de 1 a 7). Para medir a força da associação entre a percepção prévia da performance em Matemática, os componentes relativos à atitude e ao número de módulos prévios, foi utilizado o coeficiente de correlação de Spearman. A análise Estatística foi realizada com o auxílio do software IBM SPSS versão 20.

Segundo os autores, os estudantes participantes da pesquisa tendem a apreciar a utilidade e a relevância da Estatística em sua vida pessoal e profissional (Valor) e estão preparados para se esforçar a aprender Estatística (Esforço). Os alunos demonstraram uma atitude neutra a positiva quanto ao interesse em Estatística (Interesse) e à competência cognitiva (Cognitivo). Quanto à probabilidade de os estudantes realizarem

um curso de Estatística se tivessem tido a chance, 33% responderam que não o fariam (escore 1). Mesmo que 85% dos estudantes tenham tido algum curso quantitativo no passado, somente 24% foram positivos ao fato de virem a ter um (escore acima de 5). O número de cursos quantitativos prévios está positivamente relacionado aos componentes Dificuldade e Cognitivo, com uma intensidade fraca a moderada ($r_s = 0,27$ e $r_s = 0,31$, respectivamente), indicando que quanto mais cursos prévios os alunos tiveram, menor é sua percepção da dificuldade e mais positiva é sua atitude em relação aos seus conhecimentos e habilidades na aplicação da Estatística. Existiu uma forte correlação entre a percepção da performance prévia em Matemática com todos os componentes atitudinais, exceto o Esforço, indicando que para uma boa percepção da performance prévia em Matemática tem-se uma atitude positiva frente à Estatística. As mulheres tiveram um escore menor em todos os componentes, exceto o Esforço. Entretanto, nenhuma diferença significativa entre os sexos foi observada. Alunos mais velhos tiveram escores menores em todos os componentes, exceto Interesse. Ser mais velho é um preditor significativo para Dificuldade, depois do ajustamento para gênero, nacionalidade, se teve ou não curso quantitativo e performance prévia em Matemática. Ter tido ou não um curso quantitativo prévio não foi um preditor significativo para nenhum componente atitudinal. Contudo, a percepção prévia de performance em Matemática foi um forte preditor de todos os componentes atitudinais, exceto Esforço.

Hannigan et al. (2014) encerram enfatizando a diferença entre Matemática e Estatística. Eles ponderam que a Estatística não se originou da Matemática, e o papel do contexto, da variabilidade e da produção de dados, diferencia o pensamento estatístico do matemático, além disso, o pensamento estatístico depende muito da interpretação e do pensamento crítico.

Os autores ainda salientam que bons conhecimentos em Matemática não necessariamente se transformam em boa performance em Estatística, e que alunos que estudam Estatística a partir de outras áreas, como a Sociologia, por exemplo, têm atitudes mais

positivas e concepções conceituais mais adequadas dos fundamentos da Estatística do que os estudantes de Matemática. Hannigan et al. (2014) remetem, ainda, a um trabalho de Hannigan, Gill e Leavy (2013), o qual dispõe que embora os professores fossem mais capacitados matematicamente e mais confiantes, uma amostra de professores de Ensino Médio não obteve melhor desempenho em um teste avaliativo amplamente utilizado (Comprehensive assessment of outcomes in a first statistics course) do que estudantes provenientes de cursos de abordagem não quantitativa.

Para Milic et al. (2016), as atitudes subjetivas dos alunos com relação à competência cognitiva, ligada diretamente ao conhecimento matemático, afetam suas atitudes até o final do curso que, por sua vez, influenciam a sua performance. Assim, temos que a atitude negativa comumente demonstrada antes do início do curso de Estatística pode perdurar até o seu final, caso não haja um reconhecimento, por parte do aluno, de sua evolução Matemática. Também é preciso considerar, nestas atitudes, o importante fator da ansiedade, que pode estar correlacionada à ansiedade do teste em si (Hembree, 1988), exigindo um cuidado extra por parte do professor.

Como foi possível observar, a revisão de Garfield (1995) abordou a questão psicológica, porém desde um ponto de vista cognitivo restrito: raciocínios estatísticos inapropriados são generalizados e persistentes, ocorrem em todos os níveis de idade, inclusive entre pesquisadores experientes e são difíceis de mudar. A pesquisa psicológica atual avançou significativamente, incluindo os aspectos afetivos na cognição. Isto parece se refletir nos trabalhos desta revisão, pois destaca a caracterização da ansiedade estatística, que também é generalizada, persistente e de difícil superação. Ambas as abordagens psicológicas são interdependentes, pois a conjugação de esforços de alunos e professores vistos nos princípios propostos por Garfield e Ben-Zvi (2007) parece ser capaz de diminuir a ansiedade estatística. Os princípios construção do conhecimento (primeiro), envolvimento (segundo) e prática (terceiro), alinham-se com as metodologias participativas que, segundo Espindola Artola et al. (2014), são capazes de reduzir a

ansiedade estatística. Por outra parte, os princípios seis (consciência dos erros) e oito (feedbacks) expõem os alunos a uma tarefa metacognitiva de revisão de seus rendimentos, o que poderia favorecer, segundo os resultados atuais, a diminuição na ansiedade, dado que desempenho e ansiedade estão relacionados, como foi observado no trabalho de Vigil-Colet, Lorenzo-Seva e Condon (2008). Entretanto, os princípios propostos por Garfield e Ben-Zvi (2007) nada dizem a respeito da contextualização, embora este aspecto seja mencionado na sua revisão de literatura.

É interessante ressaltar que os trabalhos de Hernandez, Santos, Silva, Mendes e Ramos (2015) e Turik, Viali e Moraes (2012) apontam que estudos estatísticos contextualizados à realidade do estudante, contribuem para a diminuição da ansiedade estatística. Outra questão abordada na revisão feita por Garfield (1995), mas que não está explícita em seus princípios, é referente aos conhecimentos prévios dos alunos, principalmente os conhecimentos matemáticos que eles vêm desenvolvendo desde a escola básica, que propicia uma aproximação dos conhecimentos estatísticos e, como observam Hannigan et al. (2014), influenciam na ansiedade estatística.

De outra parte, atitudes positivas frente à Estatística estão relacionadas também às experiências prévias de sucesso em Matemática (Milic et al., 2016), daí a importância de feedbacks consistentes, tal como Garfield e Bem-Zvi (2007) expressam em seu princípio número sete, pois nem sempre o aluno é capaz de mensurar seu próprio desenvolvimento, sendo este um papel do professor.

2.6.2 Metodologias de ensino

Os artigos que trazem uma preocupação com a maneira que a Estatística é trabalhada (13) perfizeram uma quantidade parecida a dos que tratam dos aspectos psicológicos (13) e estão distribuídos como aparece na tabela abaixo. Quando os artigos descrevem um processo de ensino de Estatística efetivamente desenvolvido junto aos alunos, acrescido da análise dos resultados obtidos, esses foram classificados como de metodologia aplicada.

Caso contrário, os artigos foram classificados como contendo metodologia não aplicada, de cunho teórico. Sete trabalhos são provenientes da área da saúde e seis do ensino, destacando o esforço didático de pesquisadores que publicaram em revistas ou eventos da área da saúde na qualificação das metodologias de ensino.

Atenta-se à quantidade maior de artigos com metodologia aplicada (8) contra os de metodologia não aplicada (5). Destes, Smith (2010) destaca a importância de se diminuir a distância entre o que é necessário e o que se ensina de Bioestatística, tema que também é abordado pelo estudo dos autores Herman, Notzer, Libman, Braunstein e Steinberg (2007), ao salientarem que o currículo atual não ensina o que os médicos necessitam. Neste sentido, Macdougall (2010) enfatiza a importância da promoção da aprendizagem autônoma. Dos trabalhos que não aplicaram uma metodologia, somente Fuente e Gea (2013) analisam formas de raciocínio sobre independência.

Interpretação própria dos artigos na categoria Metodologias de ensino.

Autores	Área/ Sub- categoria	Objetivo	O que foi medido	Como foi medido	Metod. de pesquisa	Metod. de ensino	Teoria de Aprend.
(Smith, 2010)	Ensino/Teórico.	Discutir as necessidades entre pesquisadores e clínicos sobre Estatística e como melhorar o currículo	Não se aplica	Não se aplica	Discurso Teórico	Proposta de currículo	Não
(Lima, 2010)	Saúde/Aplicado	Integrar a Epidemiologia e a Estatística no ensino de Estatística	Não se aplica	Não se aplica	Quali+Relato de experiência	Desenvolveram a própria metodologia ativa mesclando epidemiologia e Estatística	Não utilizaram
(Fuente & Gea, 2013)	Ensino/Teórico	Relatar problemas dos estudantes de Psicologia com atividades de Independência	Respostas sobre correlação	Questionário	Quali+Quantitativo	Onto-semiótica	Não-utilizaram
(Turner et al., 2016)	Saúde/Aplicado	Propor uma metodologia de ensino de Estatística utilizando dados oncológicos	Não se aplica	Não se aplica	Quali+Quantitativo	Workshops mesclando problemas oncológicos com Estatística	Não utilizaram

(McGready & Brookmeyer, 2013)	Saúde/Aplicado	Verificar a diferença de rendimento de alunos da modalidade a distância e presencial	Notas	Provas finais	Quantitativa	Tradicional	Não utilizaram
(Felgueiras, 2013)	Ensino/Aplicado	Relato de experiência do uso de exemplos da área nas aulas de Estatística	Não se aplica	Não se aplica	Quali	Tradicional com uso de dados <u>não reais</u> mas familiares e exemplos das áreas dos alunos.	Não utilizaram
(Herman, Notzer, Libman, Braunstein, & Steinberg, 2007)	Saúde/Teórico	Verificar o conteúdo remanescente após as aulas de Estatística	Conhecimento remanescente após as aulas de Estatística	Questionário fechado	Quali+Quantitativa	Não se aplica	Não se aplica
(Milic, Trajkovic, et al., 2016)	Saúde/Aplicado	Comparar o rendimento de alunos expostos ao ensino híbrido com o de quem experimentou o ensino presencial	Rendimento dos alunos ao final dos dois processos	Notas ao final dos processos	Quali+Quantitativa+Inferencial+Correlação	Metodologia tradicional em ambas as modalidades.	Não utilizaram
(MacDougall, 2010)	Ensino/Teórico	Sugerir meios de se alcançar uma aprendizagem autônoma	Não se aplica	Não se aplica	Quali	Cita algumas técnicas, não metodologias	Não utilizou
(Bahassy, 2015)	Ensino/Aplicado	Comparar o rendimento de alunos expostos ao ensino convencional e os que utilizaram softwares e dados reais	Rendimento dos alunos expostos a diferentes métodos de ensino	Notas ao final do período	Quantitativa+Inferencial	Ensino tradicional baseado em software com uso de dados reais	Não utilizou
(Nowacki, 2015)	Ensino/Aplicado	Apresentar e analisar uma metodologia de ensino usando um material concreto	Não se aplica	Não se aplica	Quali	Resolução de Problemas e recomendações do GAISE	Não utilizaram

(Feild, Belgado, Dougherty, Doering, & Gong, 2015)	Saúde/Aplicado	Avaliar se a análise de literatura média melhora a confiança dos estudantes em suas habilidades.	Confiança dos estudantes em suas habilidades.	Escala própria tipo Likert	Quantitativo Inferencial	application-based teaching strategies	Não utilizará
(Nowacki, 2011)	Ensino/Aplicado	Aplicar a estrutura 4Mat no desenvolvimento da metodologia PBL e verificar a ansiedade estatística gerada no processo	Satisfação do estudante com o curso.	ATSS	Qualitativo+Quantitativo Inferencial	4Mat PBL	Construtivismo

Metodologias de ensino de Estatística inadequadas impactam na formação do futuro profissional da área da saúde, visto que a Medicina Baseada em Evidências é uma realidade e os estudantes podem estar se posicionando de forma negativa perante esta demanda, por causa da Estatística e da Matemática que ela exige. Nos Estados Unidos, foram criadas as Guidelines for Assessment and Instruction in Statistics Education (GAISE), que tiveram seu impacto também no ensino de Estatística para a área da saúde, e priorizam:

- Ensinar o raciocínio estatístico;
- Ensinar Estatística como um processo investigativo de resolução de problemas e tomada de decisão;
- Proporcionar aos estudantes experiência com o pensamento multivariado;
- Focar no entendimento dos conceitos;
- Integrar dados reais com um contexto e propósito;
- Fomentar a aprendizagem ativa;
- Utilizar a tecnologia para explorar conceitos e analisar dados;
- Utilizar as avaliações para melhorar e avaliar a aprendizagem dos alunos.

As pesquisas concentraram-se em experimentos de ensino restritos principalmente às classes dos próprios pesquisadores, refletindo concepções particulares em suas disciplinas,

ou das instituições, em seus currículos. Embora tal abordagem continue sendo maioria, a despeito dos alertas feitos por Garfield (1995), agora existem parâmetros para o ensino de Estatística, tais como as GAISE, que nortearam o trabalho de Nowacki (2015). Este autor propõe um ensino que envolva o aluno em todas as etapas da pesquisa, mediante atividades que incluam design, levantamento e coleta de dados, a fim de que os estudantes possam perceber o todo do processo de trabalho (Nowacki, 2015. p. 3).

A pesquisa de Nowacki (2015) contou com 32 estudantes de Medicina, divididos em grupos de quatro, que desenvolveram sua própria questão de pesquisa baseada em leituras prévias disponibilizadas pelo professor. Este tipo de metodologia contempla a contextualização, promove o processo investigativo e fomenta a aprendizagem ativa. Neste sentido, o trabalho de Lima (2010) mostra o resultado da integração da disciplina de Epidemiologia com a Estatística, desenvolvida entre 2002 e 2008, envolvendo 240 alunos do curso de Medicina de uma universidade brasileira. Os alunos foram incentivados a gerar e desenvolver suas próprias questões de pesquisa, a partir das necessidades da saúde da região. Como resultado extraclasse, a pesquisadora destaca a determinação do perfil epidemiológico de algumas doenças endêmicas na região.

Outro elemento das propostas metodológicas diz respeito às teorias psicológicas ou didáticas que as sustentam. Como se pode observar na Tabela iv, um único trabalho embasou-se, ao menos de forma explícita, no construtivismo, a despeito da indicação da importância de seu uso dada por Garfield (1995).

Relativamente ao uso da tecnologia, fica clara a larga utilização de computadores nas aulas de Estatística, sendo a informática uma das responsáveis pela popularização da disciplina para a área da saúde, com softwares especializados como o “Epi info”, criado pelo CDC (Center for Disease Control and Prevention) para trabalhar com dados epidemiológicos, entre outros softwares. Entretanto, sobre a acessibilidade da informática, Tishkovskaya e Lancaster (2012) pontuam a falta de estudos que incluam a tecnologia de maneira pedagógica nas aulas, e não como simples substituta da álgebra

Estatística. Contudo, a Web tem sido terreno fértil para experimentações pedagógicas diferenciadas. Prova disso é a disseminação das MOOC's (Massive Open Online Courses), disponibilizado por renomadas instituições de ensino como o MIT (Massachusetts Institute of Technology), mas ainda focadas em aulas gravadas, segundo Tavares (2014). Outrossim, nenhum dos artigos da amostra avalia alguma metodologia de ensino de Estatística de forma totalmente a distância.

Tishkovskaya e Lancaster (2012) indicam que, com relação à quantidade de opções e larga utilização da internet, é preciso uma metodologia adequada para o seu uso eficiente como ferramenta de ensino, mesmo que as propostas sejam centradas no aluno (student-centred learning) ou dirigidas por ele (self-directed learning). A intervenção do professor é requerida para um aprendizado efetivo. Propostas que conjugam o presencial com os recursos on-line tiveram um desempenho positivo.

É pertinente lembrar que Garfield (1995) já mencionava a necessidade de se desenvolver metodologias para o bom uso das tecnologias. A questão não é o uso de computadores, softwares e internet em si, mas a carência de metodologias neste contexto, que contenham um fundamento teórico consistente e aporte de pesquisa científica. Por outro lado, em seu conjunto, destacam-se as propostas que utilizam o rendimento dos estudantes como forma de medida de conhecimento, sobre os problemas desta prática apontados por Garfield e Ben-Zvi (2007).

Havendo discorrido sobre o ensino de Estatística para o Ensino Superior de maneira geral, amparado pelas revisões de Garfield (1995) e Garfield e Bem-Zvi (2007) e contando com a revisão própria de literatura sobre o ensino de Estatística, em nível superior, para a área da saúde, parte-se para as considerações finais deste trabalho.

2.7 Considerações Finais Acerca da Revisão de Literatura

Mesmo havendo algum distanciamento do foco das revisões de Garfield e a nossa revisão da literatura, percebe-se que uma quantidade dos problemas indicados por Garfield repercute no ensino de Estatística para a área da saúde. Assim, os princípios quatro (subestimar as dificuldades de entendimento) e cinco (superestimar os conhecimentos), dados por Garfield e Ben-Zvi (2007), revelam a necessidade de se desenvolver escalas para a caracterização do conhecimento dos estudantes, além das que medem a ansiedade estatística. Neste sentido, os trabalhos de cunho psicológico foram proíficos na determinação dos aspectos da ansiedade estatística, mas não foram encontrados, nestes dez anos de pesquisa, trabalhos relativos à caracterização do conhecimento estatístico dos estudantes das áreas da saúde.

No que diz respeito às metodologias de ensino, os artigos encontrados apresentam grande variedade de enfoques, abordando desde propostas curriculares a workshops, utilizando uma boa gama de propostas que mesclam a clínica médica, a pesquisa e o ensino, o uso dos computadores e da internet e, por vezes, um ambiente híbrido de ensino, na tentativa principal de atrelar a Estatística ao fazer típico do profissional da área da saúde.

Ao se observar a atuação de um médico durante a anamnese, por exemplo, percebe-se que a investigação é inerente ao seu fazer profissional. As perguntas feitas na anamnese permitem o levantamento de dados, a formulação de hipóteses, a ponderação frente à variabilidade e à incerteza e à tomada de decisão, passos do raciocínio estatístico indissociáveis da atuação do médico. No entanto, ao se adotar metodologias tradicionais de ensino de Estatística, não se transmite essa ideia, contribuindo para a ansiedade Estatística e para o distanciamento do futuro médico desta importante ferramenta.

Torres e Rossi (2013) salientam que se o profissional da área da saúde considerar a Estatística somente como uma ferramenta, correrá o risco de diminuir sua importância

e simplificar seu uso, incorrendo em erros de aplicação. Entretanto, considerando que o conhecimento estatístico está na essência da pesquisa quantitativa, simplificar seu uso ou diminuir sua importância impacta na visão científica do profissional. Por esta linha de pensamento, a qualificação dos profissionais da área da saúde pode variar de acordo com seus conhecimentos de Estatística, corroborando com a ideia de Markert (2013), que afirma que se o uso inapropriado da Estatística for evitado na divulgação científica, o ensino e a aprendizagem da disciplina aos estudantes da saúde serão melhorados e, conseqüentemente, os cuidados com a saúde dos pacientes será aperfeiçoado. Reforça-se, dessa forma, a necessidade de qualificar o ensino através de práticas que invistam no papel de investigador do aluno. No entendimento de Campos, Wodewotzki, & Jacobini (2011, p. 14),

(...) os estudantes, de um modo geral, devem ser preparados para levantar problemas de seu interesse, formular questões, propor hipóteses, coletar dados, escolher os métodos estatísticos apropriados, refletir, discutir e analisar criticamente os resultados considerando as limitações da Estatística, sobretudo no que se refere à incerteza e variabilidade.

Diante do exposto, constata-se que o uso de metodologias baseadas em teorias de aprendizagem, desenvolvidas em ambientes que valorizam a investigação sobre dados reais e amparadas por recursos computacionais desenvolvidos em ambientes híbridos (presencial e virtual), parece estar de acordo com a evolução que o ensino de Estatística demanda e que pode melhorar o aprendizado de estudantes da área da saúde.

Outro aspecto a destacar desta abordagem é que ela permite distinguir a Matemática e a Estatística. Uma tarefa importante do professor consiste no estabelecimento dos papéis de cada disciplina, evidenciando suas semelhanças e, principalmente, suas diferenças. Este intento parece ser muito difícil de ser alcançado por uma metodologia tradicional, na qual a teoria é apresentada seguida de exemplos e terminando em aplicações, pois este é o caminho que comumente se toma nas aulas de Matemática.

Em relação à Web, encontram-se diversas possibilidades novas de ensino de Estatística como o e-learning ou o b-learning ou ainda o self-learning: Porém, em sua maioria, preservam o enfoque tradicional, limitando-se os esforços ao uso desta ferramenta como dispensadora da álgebra envolvida nos cálculos estatísticos e mudando somente a modalidade de ensino. Este campo carece de pesquisas que proponham metodologias com forte embasamento pedagógico. Com vistas a superar estas dificuldades, Tishkovskaya e Lancaster (2012) ponderam sobre a necessidade de uma contínua revisão do processo de ensino-aprendizagem da Estatística e que uma reforma no ensino da disciplina e um currículo baseado numa forte sinergia entre conteúdo, pedagogia e tecnologia são necessários.

No entanto, é importante salientar que, segundo nossa revisão, os trabalhos que propõem metodologias diferenciadas ainda são, comumente, distantes das propostas sugeridas por educadores Estatísticos, tal como a GAISE, por exemplo, e dissipam-se em ações isoladas propostas normalmente por não educadores, como médicos ou estatísticos, que parecem não atentarem para as teorias de aprendizagem. Além das propostas aqui revisadas, consideramos que outra relação simbiótica da Estatística com a área da saúde e que pode gerar bons resultados pedagógicos, desde que bem empregada, é a metodologia chamada de PBL (Problem Based Learning). Esta metodologia, muito desenvolvida para o ensino da área médica, pode proporcionar o ambiente investigativo propício ao desenvolvimento pedagógico da Estatística, considerando o contexto clínico no aprendizado.

Por outro lado, chama atenção o não aparecimento de artigos que tratam de instrumentos de pesquisa para a avaliação de resultados de aprendizagem para a área da saúde. Se considerar-se que existe a preocupação com a ansiedade estatística e com as metodologias de ensino empregadas, poder-se-ia esperar pesquisadores preocupados em acessar o nível de entendimento sobre Estatística atingido pelos alunos da área da saúde. Neste sentido, cabe lembrar o que Zieffler et al. (2008) pontuam: maneiras tradicionais

de acessar o raciocínio estatístico dos alunos, como notas em testes ou mesmo exercícios, podem não corresponder ao que eles realmente sabem e que discussões em duplas ou pequenos grupos podem ser mais reveladoras dos conhecimentos dos alunos.

Dessa forma, constata-se que são necessários mais instrumentos validados que permitam acessar os conhecimentos sobre o raciocínio de conceitos estatísticos, tais como distribuição ou variabilidade. Também o uso de métodos qualitativos pode oferecer meios de desenvolver instrumentos quantitativos. A Avaliação de Raciocínio Estatístico (Statistical Reasoning Assessment) é um dos primeiros instrumentos objetivos desenvolvidos para avaliar o raciocínio estatístico dos alunos, publicado por Garfield em 1998, como descrevem Tempelaar, Gijsselaers e Loeff (2006). Esta revisão mostra que a questão colocada por Garfield (1995) sobre que tipos de procedimentos e materiais avaliativos informam melhor o professor sobre o entendimento dos alunos, ainda está aberta, principalmente no que se refere aos instrumentos quantitativos para a área da saúde, necessitando mais esforços neste sentido.

Capítulo 3

Fundamentação Teórica

Na revisão de literatura (capítulo 2), percebeu-se que aspectos psicológicos (ansiedade estatística e atitudes frente à estatística) e metodologias de ensino têm um papel relevante no ensino de estatística para a área da saúde e influenciam-se mutuamente.

Estresse, desmotivação e repulsa à estatística podem aparecer quando o aluno não vê conexão da disciplina com a área da saúde (Oliver et al., 2014; Espindola Artola et al., 2014). Espindola Artola et al. (2014) verificaram uma diminuição na ansiedade estatística, quando se utiliza uma metodologia participativa, e Hernandez et al. (2015) e Turik et al. (2012) concluíram que estudos estatísticos contextualizados à realidade do estudante contribuem para isso. Garfield e Ben-Zvi (2007) pontuam que a aprendizagem melhora se os estudantes tomam consciência e confrontam seus erros de raciocínio. Paralelamente, o bom rendimento prévio em matemática pode melhorar a atitude frente à estatística (Hannigan et al., 2014; Vigil-Colet et al., 2008).

De outra parte, Garfield e Ben-Zvi (2007) apontaram que estudantes aprendem:

1. pela construção do conhecimento;
2. pelo envolvimento ativo em atividades de aprendizagem;
3. a fazer bem somente o que praticam.

Entretanto, não houve pesquisa que se detivesse em como se dá a formação dos conceitos estatísticos a partir das adequações dos esquemas de pensamento do estudante frente às situações que caracterizam o conteúdo estatístico envolvido na área da saúde.

Por outro lado, há questões relativas às interpretações reducionistas dos Testes de Hipóteses, como a que expõe [Regnier \(2004, p. 1\)](#): “às vezes pode-se reduzir a decisão ao escolher entre duas hipóteses alternativas, mas além dessa redução simplificativa que mutila bastante a complexidade da realidade estudada, o problema é que essa decisão baseia-se em informações parciais”.

Diante dessas questões, percebe-se a oportunidade de se lançar mão de um quadro teórico que “propicie uma forte articulação entre os problemas a serem resolvidos e o conhecimento, e também entre esquemas, conceitos e símbolos ([Vergnaud, 1990, p.23](#))”, tal como a Teoria dos Campos Conceituais (TCC) de Vergnaud.

A respeito disso, esse autor esclarece o seguinte:

Ao longo da experiência, desenvolve-se um repertório de competências e de concepções. Esse repertório é relativo a campos variados: espaciais, técnicos, temporais, sociais, linguísticos, artísticos, científicos etc. Para identificar sub-repertórios, é necessário apoiar-se ao mesmo tempo nas disciplinas de referência e na psicologia, com a consciência de que cada competência é ligada às outras por filiações e por rupturas, que é preciso analisar. São, pois, conjuntos de competências e de conceitos que devem ser estudados. Um argumento essencial a favor do estudo de campos conceituais, mais que de conceitos isolados, é que um conceito ganha sentido em situações de grande variedade; que não se analisa uma situação graças a um conceito único, mas graças a um conjunto deles; que os mesmos aspectos do mesmo conceito não são adequados para tratar diferentes situações ou para diferentes procedimentos de tratamento ([Vergnaud, 2003b, p. 76](#)).

A TCC, segundo [Vergnaud \(2003b\)](#), é um quadro teórico que torna possível a integração de um ponto de vista psicológico da (i) relação entre os processos a curto prazo de aprendizado em situação e os processos a longo prazo do desenvolvimento cognitivo; (ii) da dialética entre uma visão do cognitivo em termos de competências e



Figura 3.1: Relação entre os enfoques psicológicos utilizados nesta tese.

de esquemas, de um lado, e em termos de conhecimentos e de concepções expressas, de outro lado, e (iii) do papel de mediações linguísticas e outras formas de mediação.

Convém, ao menos para esta tese, distinguir sem separar o duplo enfoque (figura 3.1) a que se refere o termo “psicológico”: os aspectos psicológicos identificados na revisão de literatura (que dizem respeito ao sentimento frente à estatística: ansiedade e atitude), e o ponto de vista de Vergnaud (a abordagem psicológica está no processo de cognição).

Retomando o papel da mediação, esse autor estabelece o lugar do professor,

Os professores desempenham um papel importante, não apenas explicando, mostrando e gerenciando a situação da sala de aula, mas também escolhendo de maneira cuidadosa e adequada as situações que tornam o conhecimento matemático significativo. (Vergnaud, 1990, p. 19)

A metodologia de ensino chamada *Problem Based Learning* (PBL), pormenorizada na seção 3.5, é uma metodologia ativa de aprendizagem compartilhada, baseada na resolução de problemas - popular em faculdades de Medicina.

A PBL pode auxiliar o professor a manejar as situações a que se refere Vergnaud; mitigar a ansiedade e a postura negativa frente à estatística; permitir que o estudante se

exercite no processo investigativo típico da área da saúde; além de ajudar na identificação dos esquemas de pensamento utilizados pelos alunos, contemplar o conteúdo como um campo conceitual, etc. Nesse sentido, a PBL pode ser fortalecida pela incorporação da TCC, resultando, quem sabe, na melhora dos resultados dos alunos.

A seguir, estrutura-se o quadro teórico deste estudo partindo-se da TCC, que é a teoria de aprendizagem sob a qual se assentam a metodologia de ensino PBL e as questões de pesquisa, tal como abordadas na metodologia de pesquisa (capítulo 4).

Em seguida, apresenta-se o conteúdo cientificamente aceito dos testes de hipóteses tal como são explorados na Instituição de Ensino Superior (IES) onde esta pesquisa foi levada a termo, considerando desde seu caráter histórico até as sutilezas de interpretação que lhes são inerentes.

Na sequência, delimita-se o campo conceitual (CC) a ser dominado num curso da área da saúde desta IES, levando-se em conta os vários conceitos que formam o CC e as estruturas típicas utilizadas pelos Alunos, como as suas dificuldades comuns, por exemplo.

Ao final, é pormenorizada a metodologia de ensino PBL, suas relações com a TCC e sua abordagem às questões levantadas na revisão de literatura.

3.1 A Teoria dos Campos Conceituais

Esta pesquisa é sobre Didática, que, segundo Vergnaud (2017, p.16), “é o estudo dos processos de transmissão e de apropriação dos conhecimentos levando em conta os conteúdos específicos que tais conhecimentos possuem”.

Vergnaud (2003b, p. 73) coloca a definição de Didática nos seguintes termos: “[...] estudo dos processos de aprendizado e de ensino relativos a um campo de conhecimento particular: de uma disciplina ou de uma profissão, por exemplo”. E segue ponderando que a Didática se apoia na Pedagogia, na Psicologia e, evidentemente, nos conjuntos

de conhecimentos cujo aprendizado é o objetivo a atingir, mas não se reduz a isso.

É inútil estudar a formação de um só conceito, de um só esquema, ou de uma só representação, uma vez que o desenvolvimento cognitivo, pelo aprendizado e pela experiência, diz respeito a uma grande variedade de situações, de esquemas, de conceitos e de formas de representação. Por essa razão é preciso estudar, em didática, o ensino e a aprendizagem de campos conceituais de certa extensão. Para analisar tais campos conceituais, é necessário apoiar-se ao mesmo tempo na epistemologia dos conhecimentos de referência e na análise dos comportamentos e das dificuldades dos alunos. A didática não pode dispensar nem uma nem outra (Vergnaud, 2003b, p. 77).

Os conteúdos específicos, no contexto deste trabalho, estão pormenorizados no subcapítulo 3.3 e formam o sistema de conceitos acerca dos Testes de Hipóteses para Média Populacional (THMP), no âmbito da disciplina de Bioestatística, ministrada na área da saúde em nível de graduação. A metodologia de ensino que serviu de base para a dinamização dos conhecimentos foi a do *Problem Based Learning* (descrita em 3.5), apoiada pela Teoria dos Campos Conceituais (TCC), descrita a seguir.

Gerárd Vergnaud absorve e aprofunda os ensinamentos de Piaget e de Vygotski em sua teoria dos Campos Conceituais. Ele salienta que Piaget entendia o conhecimento como um processo de desenvolvimento geral, biológico e social, entretanto, Piaget não estudou o desenvolvimento cognitivo dos adultos, nem os processos de aprendizagem, na escola e no trabalho, sob o ponto de vista do conteúdo conceitual específico, que não pode ser reduzido às categorias gerais de pensamento, mesmo considerando as interações entre cultura e conhecimento, como propunha Vygotski.

Moreira (2002, p. 8) resume a teoria de Vergnaud da seguinte forma:

[...] a Teoria dos Campos Conceituais é uma teoria cognitivista neopiagetiana que pretende oferecer um referencial mais frutífero do que o piagetiano ao estudo do desenvolvimento cognitivo e da aprendizagem de competências complexas, particularmente aquelas implicadas nas ciências e na técnica, levando em conta os próprios conteúdos do conhecimento e a análise conceitual de seu domínio.

Vergnaud (2017, p. 16) reforça que “para introduzir a Teoria dos Campos Conceituais, a necessidade de clareza demanda partir das perguntas que orientam a pesquisa em psicologia: a atividade e o processo”. O autor pormenoriza as questões científicas relativas à Psicologia da educação, resumindo-as em quatro questões principais:

1. Por que movimento de confronto com a realidade e de estruturação interna os conhecimentos aumentam?
2. Qual é o papel do mediador adulto? Qual é o papel do confronto? Qual é o papel da cooperação?
3. Que relações mantêm a linguagem e o pensamento?
4. De que modo a formação da personalidade, principalmente com seus aspectos afetivos, interage com a formação dos conhecimentos? (Vergnaud, 2003b, p. 64).

O autor afirma que a primeira questão encontra a resposta mais plausível em quatro grandes ideias:

- A atividade do sujeito que aprende,
- A oferta de situações favoráveis ao aprendizado,
- A mediação por parte das pessoas que o rodeiam,
- A utilização de formas linguísticas e de formas simbólicas para comunicar e representar (Vergnaud, 2003b, p. 65).

Segundo Vergnaud (2003b), foram os psicólogos, e principalmente Piaget, que demonstraram que o aprendizado não se dá simplesmente considerando o sujeito como um receptáculo ao qual bastaria demonstrar ou explicar os conhecimentos; em vez disso, existe uma grande atividade frente aos objetos que o rodeiam. Essa atividade permite que se estabeleçam relações estáveis entre as ações e os resultados e entre os objetos. O conceito de esquema designa essa atividade organizada que o sujeito desenvolve em face de uma certa classe de situações; é uma totalidade dinâmica funcional que organiza a conduta de forma invariante, não somente quanto à forma de agir mas também em conceitualizações implícitas. Vergnaud (2017, p. 19) ressalta que “expressamos nossos

conhecimentos tanto pelo que dizemos (forma predicativa) como através do que fazemos em situação (forma operatória)”.

Vergnaud (2009) expõe os princípios de sua teoria da seguinte maneira:

A teoria dos campos conceituais é uma teoria do desenvolvimento. Tem dois objetivos: (1) descrever e analisar a complexidade progressiva, a longo e médio prazo, das competências matemáticas que os alunos desenvolvem dentro e fora da escola, e (2) estabelecer melhores conexões entre a forma operacional do conhecimento, que consiste na ação no mundo físico e social, e na forma predicativa do conhecimento, que consiste nas expressões linguísticas e simbólicas desse conhecimento. Por tratar da complexidade progressiva do conhecimento, o modelo de campo conceitual também é útil para ajudar os professores a organizar situações e intervenções didáticas, dependendo tanto da epistemologia da matemática quanto da melhor compreensão do processo de conceitualização dos alunos (Vergnaud, 2009, p. 83).

Conforme Nogueira e Rezende (2014), trata-se de uma teoria psicológica relativa ao desenvolvimento cognitivo dos sujeitos, sobretudo, quando ligado à aprendizagem de competências complexas, relativas aos conteúdos e suas particularidades, ao serem abordados na escola. Desta forma, o conhecimento é a adaptação do esquema à situação.

Vergnaud (2017, p. 17) salienta que “a Teoria dos Campos Conceituais permitirá aprofundar o que é aprender e generalizar a todos os registros de atividade, desde o gesto até o raciocínio, a problemática do processo das formas de organização dessa atividade”.

Esse autor coloca que a TCC é um quadro teórico, de um ponto de vista psicológico que abarca a relação entre os processos a curto prazo de aprendizado em situação, e os processos a longo prazo relativos ao desenvolvimento cognitivo, a dialética entre uma visão do cognitivo em termos de competências e de esquemas de pensamento, de um lado, e de conhecimentos e concepções expressas, de outro, e ainda o papel das mediações linguísticas e outras formas de mediação.

Em primeira análise, compreende-se que a TCC estabelece um quadro teórico abrangente que ancora a análise do processo cognitivo no conteúdo específico, prestando aten-

ção às adaptações dos esquemas de pensamentos do estudante durante a interação em situação. Faz isto considerando não somente conteúdos, mas competências que precisam ser estudadas, desprendendo-se da ideia do ensino de conteúdos isolados.

Vergnaud (2003b) coloca as duas perspectivas de Campo Conceitual da seguinte forma:

Campo conceitual é um conjunto vasto, porém organizado, a partir de um conjunto de situações. Para fazer face a essas situações, é preciso um conjunto de esquemas de conceituações e de representações simbólicas. Em geral, a escola busca uma organização hierárquica das formas de organização da atividade. É também corrente tomar-se campo conceitual apenas como o conjunto de conceitos que permitem dar conta de uma situação ou de um conjunto de situações (Vergnaud, 2017, p. 30).

Um argumento essencial a favor do estudo de campos conceituais, mais que de conceitos isolados, é que um conceito ganha sentido em situações de grande variedade; que não se analisa uma situação graças a um conceito único, mas graças a um conjunto deles; e que os mesmos aspectos do mesmo conceito não são adequados para tratar diferentes situações ou para diferentes procedimentos de tratamento (Vergnaud, 2003b, p. 76).

Os esquemas de pensamento são maneiras de tomar e processar as informações, utilizando os invariantes operatórios, para agir sobre os objetos considerando uma intenção ou objetivo a atingir. Os esquemas se referem a classes de situações, sendo importante distingui-las. Por exemplo, as situações que dizem respeito ao levantamento de dados são distintas das relativas a sua análise. Logo, é necessária uma teoria que leve em conta as diferenças e complexidades específicas dos conceitos a serem trabalhados, como a teoria de Vergnaud, a qual argumenta a favor do estudo de campos conceituais mais que de conceitos isolados. Para analisar tais campos conceituais, conforme Vergnaud (2003b, p. 77), “é necessário apoiar-se ao mesmo tempo na epistemologia dos conhecimentos de referência e na análise dos comportamentos e das dificuldades dos alunos”.

[...] o primeiro desafio da Psicologia enquanto ciência cognitiva é colocar em evidência as formas de Conceitualização subjacentes à atividade, aquelas

que são simultaneamente fonte e produto desta atividade. É a principal razão teórica da introdução, dentro da definição de eskema, do conceito de Invariante Operatório (Vergnaud, 2017, p. 20)¹.

Vergnaud (2008, p. 20) define conceituação como a “identificação dos objetos do mundo e das suas propriedades e relações” e coloca que o esquema de pensamento abarca as ideias de repetição e de variação, no sentido de surpresa:

Encontramos essas duas ideias no conceito de “schème”-“esquema de pensamento”²-porque esse contém ao mesmo tempo a ideia de familiaridade (a experiência, até mesmo aquela dos grandes especialistas, se constitui em grande parte das situações encontradas frequentemente), e a ideia de novidade (a possibilidade de enfrentar novas situações). De fato, diante de uma nova situação, não estamos completamente empobrecidos. Nós recorremos aos meios do nosso repertório de “esquemas de pensamento”, nos ajudando para tudo, nos enganando muitas vezes, produzindo erros que justamente resultam da inadequação dos esquemas empregados. Depois, descobrindo novos aspectos, descombinando e recombinaando elementos de esquemas anteriores, chegamos a uma solução, pelo menos provisória, mais ou menos pertinente e local. [...] Os “esquemas de pensamento” mostram bem os dois aspectos fundamentais do pensamento: sistemático e oportunista (Vergnaud, 2008, pp. 14, 15).

A TCC utiliza duas definições de esquema de pensamento, de acordo com Vergnaud (2017)³:

Definição 1: o esquema de pensamento é uma organização invariante da atividade para uma dada classe de situações.

Definição 2: o esquema de pensamento é formado necessariamente por 4 componentes:

1. Uma meta, sub-metas e antecipações;
2. Regras de ação, de busca de informação e de controle;

¹A palavra eskema costuma ser usada no sentido de esquema de pensamento, em contextos como esse.

²Anteriormente, referido como “eskema”.

³A palavra eskema foi trocada por esquema de pensamento.

3. Invariantes operatórios: conceitos em ação e teoremas em ação⁴;
4. Possibilidades de inferência em situação.

Da primeira definição, Vergnaud depreende que o esquema de pensamento é direcionado a uma classe de situações, o que permite associar a ele quantificadores universais, capazes de definir seu alcance e seus limites. Ele é universal da mesma maneira que o conceito. O autor ressalta que as competências são compostas por esquemas de pensamento destinados a enfrentar situações, os quais não são feitos de textos, são a parte operacional do conhecimento, conforme Vergnaud (1998).

Os invariantes operatórios são compostos de conceitos e teoremas em ação. Um conceito em ação é um objeto, predicado ou categoria de pensamento considerado relevante à situação e permite identificar elementos conhecidos nas situações, enquanto que um teorema em ação é uma proposição que se supõe verdadeira sobre a situação, fornecendo as regras que vinculam tais elementos, permitindo resolver uma dada situação e agir em consequência. Este conhecimento para a ação costuma permanecer totalmente implícito, dirigido ao saber fazer, em vez do desenvolvimento de conceitualizações, por isso não são verdadeiros conceitos e teoremas os quais necessitam ser explícitos para que se possa discuti-los, de acordo com Greca e Moreira (2003).

O que é invariante é a organização e não o comportamento observável. Os comportamentos são diferentes pelas variáveis de cada situação, por isso não são estereotipados. Além disso, os esquemas de pensamento organizam a atividade de pensamento subjacente ao comportamento.

Para Vergnaud (1998), os Invariantes Operatórios são a chave para a conexão entre a teoria e a prática. Eles são cruciais pois prática é ação, e esta é sempre eficiente sob certas condições e é dirigida por regras do tipo “se...então” bem como porque haveria a possibilidade de tais regras emergirem sem serem entendidas, caso não houvesse

⁴Suprimiram-se os hifens das expressões conceitos-em-ação e teoremas-em-ação.

categorias cognitivas para analisá-las, investigar os componentes da ação e a relação entre as metas, condições e ações.

Relativamente à segunda definição, Vergnaud (ibid.) reforça seu caráter analítico, salientando que as regras de ação, de busca de informação e de controle asseguram a função geradora do esquema de pensamento, sendo aquela que é mais imediatamente responsável temporalmente da conduta e da atividade, é a parte dos esquemas de pensamento que gera o comportamento em função de variáveis das situações. Esse autor não limita a conduta somente às ações, mas também pela busca de informação e controle para a continuação apropriada da atividade, dentro do que foi previamente planejado para ser feito (metas). Tais metas, que são a parte intencional do esquema de pensamento (mesmo que não sejam totalmente conscientes), são totalmente condicionadas pela representação da meta a alcançar e pelas conceitualizações que permitem a identificação, as propriedades e as relações que a conduta do sujeito pressupõe gerar nos objetos presentes. A meta é essencial na organização da atividade e se divide em submetas, ordenadas sequencial e hierarquicamente originando numerosas antecipações.

Vergnaud (ibid.) avança realçando o papel dos invariantes operatórios:

Mais decisivos ainda do ponto de vista cognitivo, são os Invariantes Operatórios, já que os conceitos em ação permitem captar as informações pertinentes no ambiente e selecionar os teoremas em ação necessários para avaliação tanto das metas e submetas suscetíveis de formar-se, como das regras de ação, de busca de informação e de controle que possibilitam alcançá-las (Vergnaud, 2017, p. 34).

Vergnaud estabelece uma sequência de utilização dos Invariantes Operatórios, em que primeiro entram em cena os conceitos em ação, capazes de avaliar e escolher o que parece pertinente, inclusive os teoremas em ação que serão utilizados para cumprir as metas.

Os conceitos em ação permitem identificar objetos, propriedades e relações. Os objetos podem ser materiais ou construídos pela cultura, ciência, técnica ou pelo próprio

sujeito. Já as propriedades e relações são os predicados observáveis e os inferidos pelos observáveis, resultantes de elaborações culturais ou pessoais. Os teoremas em ação são proposições tidas como verdadeiras na realidade, podem ser válidas para o aqui e agora ou para universalmente verdadeiras. Os conhecimentos em ação (Invariantes Operatórios) não são necessariamente explícitos, nem explicáveis e nem sempre conscientes (ibid.).

O desenvolvimento das competências e dos conceitos são caros para Vergnaud, pois ele considera um conceito como produto da experiência, ampliando-o como resultado de uma definição, mesmo que as definições sejam importantes para a racionalidade. Entretanto, Vergnaud (1998) salienta que, embora os Invariantes Operatórios sejam a fonte para os conceitos e proposições, seu alcance de disponibilidade e validade é usualmente muito limitado e local, enquanto que os conceitos científicos têm um alcance muito maior e são organizados em sistemas integrativos. Se a ciência não é sistêmica e cumulativa, não é ciência.

Neste sentido, Vergnaud define um conceito com um triplete de conjuntos distintos, não independentes entre si, porém, diferentes:

$$\text{Conceito} = DEF(S, I, R)$$

Onde,

S : é o conjunto das situações que dão sentido ao conceito;

I : é o conjunto dos Invariantes Operatórios;

R : é o conjunto das representações linguísticas e simbólicas.

As situações e os Invariantes Operatórios já foram abordados anteriormente. Quanto as representações linguísticas e simbólicas (algébricas, gráficas, etc), elas permitem

representar os conceitos e suas relações e, conseqüentemente, as situações e os esquemas de pensamento que evocam.

Moreira (2002) salienta que o conjunto S das situações é o referente do conceito, o conjunto I dos Invariantes Operatórios é o significado do conceito e o conjunto R das representações simbólicas é o significante. Em termos psicológicos, S é a realidade e (I,R) a representação que pode ser considerada como dois aspectos interagentes do pensamento, o significado I e o significante R .

Os conceitos são formados a partir das regularidades observadas pelo sujeito, em sua ação em situação, chegando a construções conceituais de alto nível que já não têm mais relação facilmente identificável com as regularidades do real. Se não fosse a linguagem e os simbolismos desenvolvidos pela cultura, seria impossível identificar essas construções conceituais.

Assim, um Campo Conceitual é, ao mesmo tempo, um conjunto de situações e um de conceitos. Para que haja um domínio progressivo daquele, é preciso um sistema de conceitos, esquemas de pensamento e de representações simbólicas estreitamente conectados. Os sistemas de conceitos possuem uma organização progressiva, eventualmente nunca concluída.

Não obstante, está a importância da linguagem no processo de conceitualização. Vergnaud destaca que os conceitos que não são construídos a partir da identificação de padrões, nas diferentes situações, precisam da linguagem para serem conceitualizados, tal como o conceito de número, que não existe fisicamente como objeto material e precisa do estabelecimento de relações para a sua construção, as quais seriam impossíveis sem a linguagem.

3.1.1 A TCC como guia para a organização do ensino

O “esquema de pensamento” não tem somente uma função de ação sobre o real, mas também uma função de interrogação do real (Vergnaud, 2008, p. 20). Daí a importância

de se proporcionar um ambiente de pesquisa durante o ensino de estatística, visto que as situações típicas do processo investigativo, cedo ou tarde, exigirão a desacomodação dos esquemas dos estudantes frente aos novos desafios. Além disso, [Vergnaud \(2017\)](#) salienta que não se pode opor a experiência cotidiana do sujeito das experiências escolares, profissionais ou de formação, pois seus esquemas de pensamento são bastante semelhantes frente a estas situações, assim, é notório que a aproximação entre o cotidiano e o científico por meio da conduta investigativa (que existe em ambos) premia os esquemas de pensamento já formados, facilitando a familiarização do sujeito frente às novas situações.

Portanto, durante o processo investigativo, o professor tem a chance de levantar indícios dos invariantes operatórios utilizados por cada um dos alunos, precisando lançar mão de instrumentos variados para tal, visto que não é uma tarefa fácil. Mesmo as questões típicas apresentadas em livros textos de estatística podem apresentar tais indícios. É preciso levar em conta, como destaca [Vergnaud \(2017\)](#), que a ação de mediação típica do professor encontram apoio na definição de esquema de pensamento pois ajudam na identificação da meta a alcançar, nas sub-metas e antecipações.

[Vergnaud \(2017, p. 15\)](#) postula que “O conhecimento é adaptação”, no sentido que adaptação é a modificação proveniente da vivência em situação. O indivíduo se adapta às situações por meio dos seus esquemas de pensamento que se modificam ao enfrentarem novas situações.

A necessidade da evolução das situações quanto a sua complexidade na busca do desenvolvimento dos conhecimentos do sujeito, coloca em evidência a ação mediadora do responsável pela sua educação. Porém, a ação do mediador não se reduz a selecionar situações potenciais promotoras da atividade do sujeito, mas também de lhe dar modelos e indicar caminhos, incitando-o a ação [Vergnaud \(2003b, 2017\)](#). Neste sentido, Vygotski definiu a zona de desenvolvimento proximal como sendo o que o sujeito consegue fazer com a ajuda dos outros e que não conseguiria sozinho [Vergnaud \(2003b\)](#). Daí se infere a

importância do debate, da discussão, da interação entre os atores do processo de ensino, para que a atividade do sujeito que aprende possa manifestar-se verbal e simbolicamente.

Além disso, o nível dos conhecimentos são modificados pelos significantes linguísticos e formas simbólicas utilizadas no ensino durante a ação em situação, daí a importância da linguagem, e de suas relações com o pensamento. A atividade linguística do professor reflete a organização dos esquemas e ajuda os estudantes a formar novos esquemas em situação [Vergnaud \(2003b\)](#). Entretanto, a ação externa da linguagem pelo mediador é insuficiente para a conceitualização visto esta é um processo interno de construção frente a experiência, que necessita do auxílio do meio, ou seja, da qualidade da mediação.

Relativamente à análise das relações entre linguagem e pensamento, [Vergnaud \(2003b\)](#) destaca a estrutura predicativa da linguagem, em particular a contida nos enunciados e textos científicos, é posta em relação com os invariantes operatórios dos esquemas dos sujeitos, transformando-os pela enunciação em teoremas explícitos. Neste sentido, a linguagem que acompanha o pensamento auxilia para a seleção e transformação da informação, na regulação e planejamento da ação e para o esclarecimento e lembrança do objetivo a atingir.

De outra parte, considerando que os esquemas se referem a classes de situações, é importante diferencia-las, por exemplo, o enfrentamento de situações referentes à álgebra são diferentes dos referentes à aritmética. A este respeito, é necessário considerar uma teoria que leve em conta as diferenças e complexidades específicas dos conceitos a serem trabalhados, como a teoria de Vergnaud, a qual argumenta a favor do estudo de campos conceituais mais que de conceitos isolados. Para analisar tais campos conceituais, é necessário apoiar-se ao mesmo tempo na epistemologia dos conhecimentos de referência e na análise dos comportamentos e das dificuldades dos alunos ([Vergnaud, 2003b](#), p. 77).

O agir pedagógico consiste em harmonizar as situações aos esquemas do sujeito de tal maneira que ocorra uma desacomodação dos Invariantes Operatórios que ele lança mão

naquela determinada situação, provocando-o a construir novos Invariantes Operatórios, ou a adaptar os antigos, caracterizando a evolução do conhecimento. Vergnaud coloca a importância desta ação mediadora do professor da seguinte forma:

Em que consiste a revolução didática? Consiste em propor ao aluno situações que vão desestabilizá-lo. Essas situações desestabilizadoras, graças à ação auxiliar do professor, poderão ser incorporadas pelo aluno, para seu proveito (Vergnaud, 2003b, p. 38).

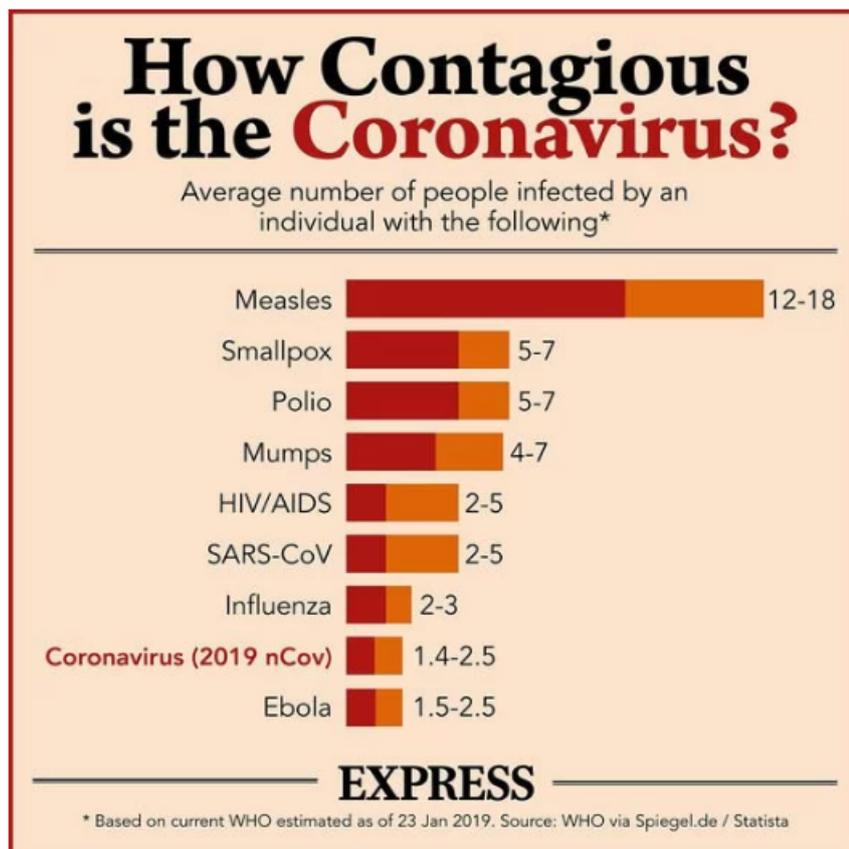
Outro ponto crucial da teoria de Vergnaud que, ao mesmo tempo que se distancia de Piaget, auxilia decisivamente no fazer didático do professor é o distanciamento dos modelos teóricos da lógica em benefício do conteúdo conceitual dos conhecimentos. Desta forma, a visão piagetiana de abordagem da inteligência, basicamente, em termos lógicos, é substituída em Vergnaud, pelos processos de conceitualização que abarcam todos os registros da atividade (ibid.).

Se por um lado a teoria de Vergnaud considera campos conceituais em vez de conceitos isolados e sustenta que são necessárias diferentes situações para dar sentido aos conceitos, então, a conceitualização, aceita cientificamente, em certo ponto, é fruto histórico das interações dos invariantes operatórios explícitos daqueles que contribuíram para a sua construção. Neste sentido, a ciência é o resultado aceito temporalmente, dos conceitos e teoremas em ação daqueles capazes de explicitá-los e confirmá-los de maneira consistente para manterem-se válidos após o julgamento dos seus pares no decorrer do tempo. Desta forma, o conhecimento cientificamente aceito constitui-se o padrão ouro desejável que o estudante domine, sendo possível “medir” o seu conhecimento através da análise da distância entre os conceitos científicos explícitos e dos invariantes operatórios explicitados pelo sujeito aprendente. Assim, é preciso colocar de forma explícita, da melhor maneira possível, o campo conceitual que se quer trabalhar, o que para o caso dos testes de hipóteses em geral e para uma média são expostos a seguir.

3.2 Raciocínio, Letramento e Pensamento Estatístico

Qual a importância da estatística para a área da saúde?

Durante o surto de corona vírus de 2020, a estatística foi utilizada para informar os leitores do jornal Daily Express, como mostra a figura 3.2.



UK coronavirus: The new disease has a smaller mortality rate than SARS (Image: EXPRESS)

Figura 3.2: Imagem veiculada na matéria “Taxa de mortalidade por coronavírus no Reino Unido: o coronavírus matou mais do que a SARS - Qual é o perigo?” (Kettley, 2020)

A figura 3.2 insere-se no contexto da epidemiologia, que para Bonita et al. (2010, p. 1), “[...] é uma disciplina relativamente nova e usa métodos quantitativos para estudar a ocorrência de doenças nas populações humanas e para definir estratégias de prevenção e controle”.

Outro enfoque que também envolve métodos quantitativos na área da saúde trata da medicina baseada em evidências (MBE).

A medicina baseada em evidências (MBE) é um novo paradigma para o sistema de assistência médica, envolvendo o uso das evidências atuais (resultados de estudos de pesquisas médicas) na literatura médica para fornecer o melhor atendimento possível aos pacientes (Mayer, 2009, p. 1).

A MBE, que para Duncan et al. (2014, p. 44), seus princípios, regras e informações de apoio à decisão clínica são amplamente aplicáveis às outras áreas clínicas da saúde, pode ser vista da seguinte forma.

Fundamenta-se na aplicação de conhecimentos básicos de epidemiologia e bioestatística para avaliar a evidência clínica quanto a sua validade e utilidade potencial. Praticar com base em evidências é integrar as melhores evidências de pesquisa à habilidade clínica do profissional e à preferência do paciente (da Cruz e Pimenta, 2005, p. 416).

Parece que enquanto a epidemiologia trata das populações, a MBE ressalta o tratamento ao paciente particular e ambas lançam mão da estatística.

Neste cenário, o conhecimento estatístico do profissional da área da saúde relaciona-se ao seu fazer clínico, conquanto o conhecimento estatístico do paciente refere-se ao entendimento do seu estado de saúde.

Para ilustrar, a figura 3.3, mostra um resultado de exame, disponível tanto ao profissional da área da saúde quanto ao paciente.

O exame da figura 3.3 apresentou o resultado mostrado na figura 3.4.

A interpretação do resultado do exame, tomando-se sua forma mais simples possível, compara o Escore T do paciente com a tabela da OMS, caracterizando osteoporose na coluna lombar e osteopenia no colo do fêmur. Esta interpretação não requer conhecimento estatístico, mas não permite inferir o grau destas enfermidades.

Um entendimento mais abrangente do exame (figura 3.3) exige mais do que a alfabetização funcional, é preciso uma literacia estatística.

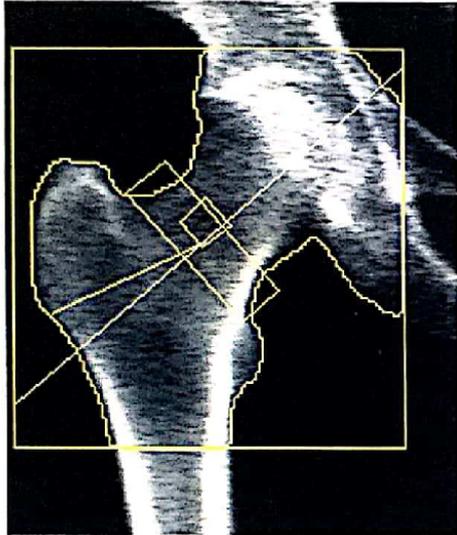


Imagem não destinada a diagnóstico
 k = 1.136, d0 = 46.1
 109 x 113
 PESCOÇO: 49 x 15

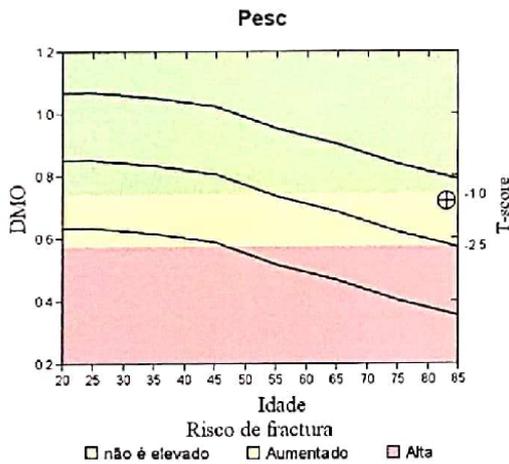
Informações do exame:

Data do exame: 29 Novembro 2019 ID: A1129190M
 Tipo de Exame: F Fémur direit
 Análise: 29 Novembro 2019 17:01 Versão 13.5.2.1:5
 Fémur
 Operador: SLC
 Modelo: Discovery Wi (S/N 88359)
 Comentário:

Resumo dos resultados de DXA:

Região	Área (cm²)	BMC (g)	DMO (g/cm³)	T - pont.	PR (%)	Z - pont.	AM (%)
Pesc	5.62	4.03	0.717	-1.2	85	1.3	124
Troc	12.00	8.21	0.684	-0.2	97	1.7	134
Inter	19.41	17.31	0.891	-1.3	81	0.6	112
Total	37.04	29.55	0.798	-1.2	85	1.0	119
Ward	1.06	0.51	0.476	-2.2	65	0.9	129

DMO total CV 1.0%, ACF = 1.012, BCF = 0.988, TH = 5.750



Comentário:

Pontuação T comparação com Branco Mulher; Pontuação Z Branco Mulher.
 Origem: BMDCS/NHANES White Female.

HOLOGIK

Figura 3.3: Parte de um exame densitometria óssea digital de uma paciente de 83 anos de idade.

Data	29/11/2019	Sexo	Feminino	Idade	83a, 0m, 15d
COLUNA LOMBAR E FÊMUR PROXIMAL					
REGIÃO	DENSIDADE MINERAL ÓSSEA (BMD – g/cm ²)	ADULTO JOVEM		CORRIGIDO PARA IDADE	
		Escore T	%	Escore Z	%
LOMBAR (TOTAL)	0.770	-2.5	74	0.3	104
FÊMUR (PESC)	0.717	-1.2	85	1.3	124
COMENTÁRIO					
<p>COLUNA AP : A densidade mineral óssea das vértebras lombares L1-L4 é de 0.770 g/cm². Este valor está 2.5 DP abaixo ao da população controle jovem.</p> <p>COLO DO FÊMUR : A densidade mineral óssea do fêmur proximal é de 0.717 g/cm². Este valor está 1.2 DP abaixo ao da população controle jovem.</p>					
CRITÉRIOS DENSITOMÉTRICOS DA ORGANIZAÇÃO MUNDIAL DA SAÚDE (OMS)					
CLASSIFICAÇÃO	T-ESCORE				
NORMAL	-1.0 OU MAIOR				
OSTEOPENIA	ENTRE -1.0 E - 2.5				
OSTEOPOROSE	-2.5 OU MENOR				
Comentário					
<p>COLUNA LOMBAR: Há critérios para OSTEOPOROSE. COLO DO FÊMUR : Há critérios para OSTEOPENIA.</p>					

Figura 3.4: Resultado do exame densitometria óssea digital da 3.3.

A literacia estatística inclui habilidades básicas e importantes que podem ser usadas para entender informações estatísticas ou resultados de pesquisas. Essas habilidades incluem a capacidade de organizar dados, construir e exibir tabelas e trabalhar com diferentes representações de dados. A literacia estatística também inclui uma compreensão de conceitos, vocabulário e símbolos e inclui uma compreensão da probabilidade como uma medida de incerteza (Ben-Zvi e Garfield, 2004, p. 7).

No caso do gráfico na figura 3.2, pouco adianta o conhecimento da língua inglesa na interpretação da figura, senão, é necessário conhecimentos acerca de média, gráficos, proporções e outros conceitos estatísticos, para que haja um entendimento sobre o que os números, desenhos e cores da figura querem dizer.

Tal conhecimento precisa estar organizado em esquemas de pensamento para darem conta do entendimento desta situação, logo, a literacia estatística necessária nesta situação conta com conceitos dispostos em um campo conceitual, que, segundo Vergnaud (2003a, p. 31), é um “[...] conjunto de conceitos que permitem dar conta de uma situação ou de um conjunto de situações”.

Esta estrutura organizada de conceitos chamada de campo conceitual da estatística refere-se ao conhecimento científico aceito e explicitado e passível de ser dominado em diferentes graus, da literacia ao pensamento estatístico, passando pelo raciocínio estatísticos, tal como Ben-Zvi e Garfield (2004, p. 7) os definem.

O raciocínio estatístico pode ser definido como a maneira como as pessoas raciocinam com estatísticas e compreendem as informações estatísticas. Isso envolve fazer interpretações baseadas em conjuntos de dados, representações de dados ou resumos estatísticos de dados. O raciocínio estatístico pode envolver a conexão de um conceito a outro (por exemplo, central e dispersão), ou pode combinar ideias sobre dados e chance. Raciocinar significa entender e ser capaz de explicar processos estatísticos e interpretar completamente os resultados estatísticos (Ben-Zvi e Garfield, 2004, p. 7).

O pensamento estatístico envolve uma compreensão de por que e como as investigações estatísticas são conduzidas e as “grandes ideias” subjacentes às investigações estatísticas. Essas ideias incluem a natureza onipresente da

variação e quando e como usar métodos apropriados de análise de dados, como resumos numéricos e exibição visual de dados. O pensamento estatístico envolve uma compreensão da natureza da amostragem, como fazemos inferências de amostras para populações e por que experimentos projetados são necessários para estabelecer a causalidade. Inclui uma compreensão de como os modelos são usados para simular fenômenos aleatórios, como os dados são produzidos para estimar probabilidades e como, quando e por que as ferramentas inferenciais existentes podem ser usadas para auxiliar um processo investigativo. O pensamento estatístico também inclui ser capaz de entender e utilizar o contexto de um problema na formação de investigações e tirar conclusões, além de reconhecer e entender todo o processo (da pergunta ao levantamento de dados, à escolha de análises e testes, suposições etc.). Finalmente, pensadores estatísticos são capazes de criticar e avaliar os resultados de um problema resolvido ou de um estudo estatístico.

As ideias acima podem fazer pensar em um nível crescente de domínio da estatística, desde uma leitura interpretativa na literacia até um comportamento semelhante ao do estatístico no pensamento estatístico. Entretanto, [Delmas \(2002\)](#) aponta que há sobreposições entre elas, de tal forma que uma situação pode exigir as três competências. O autor (Op. cit.) ainda sugere que a sobreposição se dá como na figura 3.5.

No caso do exame da figura 3.3, o resultado é atribuído em termos da quantidade de desvios padrão de afastamento da média da densidade mineral óssea (DMO) de uma população de controle jovem (Escore T), apresentado por [Defavori e Sarriés \(2007, p. 186\)](#) da seguinte forma.

$$\text{Escore T} = \frac{\text{DMO medida} - \text{DMO média em adultos jovens}}{\text{SD adultos jovens}}$$

Ao considerar-se uma distribuição pelo menos aproximadamente normal da DMO, o raciocínio estatístico leva a concluir que um Escore T de -2,5 significa uma densidade mineral óssea muito inferior à média do adulto jovem, limítrofe dos 0,62%⁵ das menores DMO desta população de referência.

Por sua vez, a competência do pensamento estatístico pode vir a questionar o motivo

⁵Cálculo feito pelo *applet* disponível em <http://www.rossmanchance.com/applets/NormCalc.html>

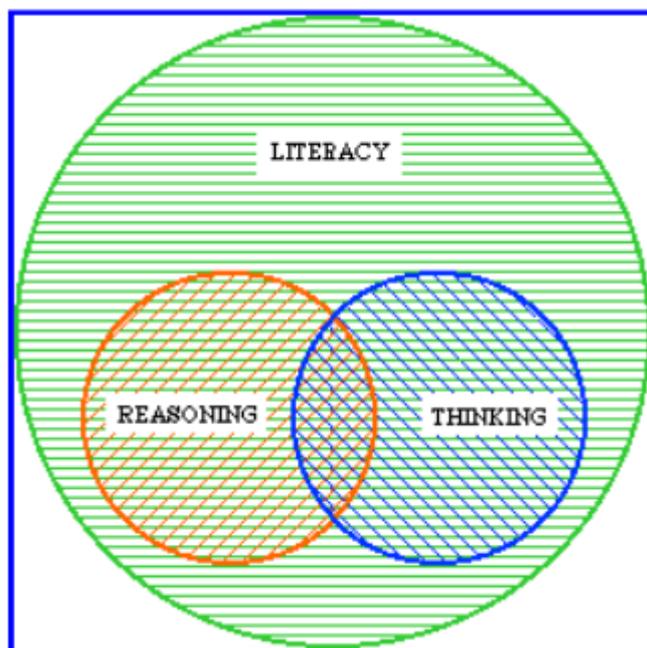


Figura 3.5: Relação entre literacia (LITERACY), raciocínio (REASONING) e pensamento (THINKING) estatístico dada por Delmas (2002, p. 5)

da comparação da média de DMO de uma idosa ser feita com a de jovens adultos e não com os de idades compatíveis dado que o Escore Z fornece esta comparação.

Ainda, o aluno conhecedor dos Escores Z e T, tal como foram postos na seção 5.2, percebe que, no contexto da densitometria óssea, tais escores diferem dos da dispersão. Neste sentido, o campo conceitual da dispersão (seção 5.2) entrelaça-se com o da densitometria óssea, mas se mantém diferentes. Neste caso, os esquemas de pensamento de um precisam adaptar-se ao outro, mas sempre estarão no contexto da literacia, raciocínio e pensamento estatístico.

As diferenças entre os escores acontece porque, embora os significantes sejam os mesmos (Escore Z e T), os significados são distintos, mesmo que ambos envolvam o conceito de variabilidade em suas definições.

O posicionamento investigativo e crítico coloca os esquemas de pensamento à prova diante das situações, o que pode ser estimulado nos Alunos por uma metodologia de

ensino que aprofunde a literacia estatística em direção do pensamento estatístico, como o PBL, por exemplo.

3.3 O Conhecimento Científico dos Testes de Hipóteses

Delinear o conhecimento cientificamente estabelecido temporalmente é central na teoria de Vergnaud, principalmente por explicitar o engendramento de diversos conceitos que se deve levar em conta ao se considerar as situações que lhes encerram.

Vergnaud coloca a questão temporal e relativa da matemática da seguinte forma,

A matemática é uma “construção social falibilista” (Ernest, 1985; Lerman, 1987), no sentido de que os conceitos se desenvolvem através da negociação com situações e com outras pessoas; conceitos são culturalmente e temporalmente relativos e potencialmente falíveis. A objetividade aparece na natureza pública compartilhada de teorias e conceitos, e não através da correspondência completa com o mundo real. Verdade, prova e rigor também podem ser vistos como relativos. (Vergnaud, 1990, p. 29)

Considerar a subjetividade dos conceitos lhes tira a impressão de verdades absolutas e inquestionáveis o que impacta em sua didática, visto que o estudante pode e deve colocar-se criticamente frente ao saber científico. Desta maneira, o delineamento do saber científico em estudo impactará na escolha das situações que ao mesmo tempo lhe dão sentido e chamam o estudante à pesquisa.

Vergnaud (2003b) pondera que a questão fundamental em didática é a escolha de situações apropriadas para os Alunos, considerando-se o desenvolvimento que já atingiram, e que esta escolha depende da epistemologia do conhecimento específico dos diversos domínios do conhecimento. Vergnaud (Op. cit.) refere-se à epistemologia como o conteúdo de conhecimento em relação àquelas questões para as quais esse conteúdo de conhecimento propõe respostas.

Esta seção descreve o conhecimento científico dos Testes de Hipóteses desde sua concepção histórica até suas sutilezas de interpretação, concentrando-se na descrição

pormenorizada dos Testes de Hipóteses para Média Populacional (THMP) tipicamente abordada em cursos da área da saúde.

Tal enfoque subsidia o delineamento do Campo Conceitual (CC) dos THMP no que diz respeito ao sistema de conceitos cientificamente aceitos que integram tal CC.

3.3.1 Estatística, Inferência Estatística e Testes de Hipóteses: uma Abordagem Típica da Área da Saúde

Vergnaud (2017, p. 18) enfatiza que “É preciso conceitos que sejam os mais bem definidos possíveis para avançar. Em particular, para conduzir pesquisas empíricas e para interpretar observações. Assegurando um lugar central à dupla teórica situação/eskema⁶”.

Vergnaud (2009) afirma que “Ciência é redução. Mas nem todas as reduções são frutíferas”. Esta seção procura definir um aparte dos sistemas de conceitos cientificamente aceitos dos Testes de Hipóteses que sejam necessários e suficientes para o estudante da área da saúde.

O conhecimento científico aqui descrito procura levar em conta as limitações típicas de estudantes da área da Saúde (desconhecimento do Cálculo Diferencial e Integral, baixa carga horária da disciplina de Bioestatística, etc.) ao mesmo tempo que tenta ser o mais abrangente possível. Tomou-se por base os livros de estatística dirigidos para a área da saúde dos autores Callegari-Jacques (2003); Chernick (2011); Pagano e Gauvreau (2006); Díaz e López (2012); Martinez (2015); Callegari-Jacques (2003) e os livros de estatística geral dos autores Morettin e Bussab (2017); Casella e Berger (2011).

A Estatística pode ser definida como segue:

Estatística é a ciência que tem por objetivo orientar a coleta, o resumo, a apresentação, a análise e a interpretação de dados. Podem ser identificadas duas grandes áreas de atuação desta ciência: a estatística descritiva,

⁶A palavra “eskema” aparece em algumas publicações de Vergnaud em português e em outras vê-se “esquema de pensamento”, em contextos similares, sugerindo serem equivalentes.

envolvida com o resumo e a apresentação dos dados, e a estatística inferencial, que ajuda a concluir sobre conjuntos maiores de dados (populações) quando apenas partes desses conjuntos (as amostras) foram estudadas. Os métodos da estatística inferencial são ferramenta imprescindível no teste das hipóteses científicas. (Callegari-Jacques, 2003, p. 14)

Outro aspecto importante da Estatística refere-se à variabilidade. Conforme Hogg et al. (2013), “A variabilidade é um fato da vida e métodos estatísticos apropriados podem nos ajudar a entender os dados coletados sob a variabilidade inerente. Devido a essa variabilidade, muitas decisões precisam ser tomadas, envolvendo incertezas.”

Neste sentido, a seguinte definição ganha abrangência:

A Estatística ocupa-se tanto dos métodos e procedimentos, para coletar, classificar, resumir, achar regularidades e analisar os dados, sempre e quando a variação e a incerteza forem causa intrínseca destes, como de realizar inferências a partir deles, com a finalidade de auxiliar na tomada de decisões e, neste caso, formular previsões. (Díaz e López, 2012, p. 4)

Para Callegari-Jacques (2003), dados são as informações, que podem ser numéricas ou não, obtidas a partir de elementos ou indivíduos que constituem a menor unidade a fornecer tal informação.

Cada dado assume um valor único dentre as possibilidades de um determinado conjunto e as características desse conjunto auxiliam na determinação do tipo de variação observada. Assim, uma variável pode assumir qualquer valor de um conjunto determinado, chamado domínio. Em função do tipo de domínio as variáveis tem a seguinte classificação, dada por Callegari-Jacques (2003):

Variáveis Qualitativas (ou Categóricas): fornecem dados de natureza não-numérica.

Nível nominal: uma categoria (dado não numérico) diferencia-se da outra somente por sua denominação.

Nível ordinal: é possível identificar diferentes categorias e também reconhecer graus de intensidade entre elas, permitindo uma ordenação das categorias. É

necessário, no entanto, que a gradação seja inerente à variável e não imposta por conveniência pelo pesquisador.

Variáveis Quantitativas: os dados são valores numéricos que expressam quantidades.

Variáveis quantitativas discretas: os dados só podem apresentar determinados valores, em geral, números inteiros.

Variáveis quantitativas contínuas: os dados podem apresentar qualquer valor dentro de um intervalo de variação possível.

As variáveis, referem-se à População, que para [Martinez \(2015\)](#), são o conjunto de todas as pessoas, sujeitos, organismos ou objetos que possuem uma ou mais características em comum e a sua análise visa responder à pergunta formulada e a ela direcionada. Já uma amostra, é uma parte selecionada da totalidade de elementos da população cuja característica importante é a de ser representativa desta população. Aqui assume-se que as amostras são coletadas adequadamente.

Considerando que populações podem ser grandes demais, ou de alguma forma inacessíveis em sua totalidade, é útil pensar na possibilidade de se concluir algo relativo à população tomando por base uma amostra desta população.

Conforme [Díaz e López \(2012\)](#), a tarefa fundamental da estatística inferencial é fazer inferências acerca da população a partir de uma amostra dela extraída.

Para estudantes de área da saúde em nível de graduação, [Díaz e López \(2012\)](#); [Martinez \(2015\)](#); [Callegari-Jacques \(2003\)](#); [Chernick \(2011\)](#) consideram em seus livros, dois ramos da inferência estatística: a Estimação e o Teste de Hipóteses.

Na visão de [Costa \(2012\)](#),

Quando quisermos avaliar um parâmetro populacional, sobre o qual não possuímos nenhuma informação com respeito a seu valor, não resta alternativa a não ser estimá-lo através do intervalo de confiança. No entanto, se tivermos alguma informação com respeito ao valor do parâmetro que



Figura 3.6: Relação entre População, Amostra e Inferência.

desejamos avaliar, podemos testar esta informação no sentido de aceitá-la como verdadeira ou rejeitá-la. [Costa \(2012\)](#)

Ainda para [Costa \(2012\)](#), a Inferência Estatística “é a parte mais importante de um processo inferencial. Todo estudo com levantamento por amostragem que mereça crédito deve realizar testes de significância de estimativas geradas”.

Para [Morettin e Bussab \(2017, p. 330\)](#), “o objetivo do teste estatístico de hipóteses é fornecer uma metodologia que nos permita verificar se os dados amostrais trazem evidências que apoiem ou não uma hipótese (estatística) formulada”. [Casella e Berger \(2011\)](#) indica que uma hipótese estatística é uma afirmação sobre um parâmetro populacional.

Ainda para [Casella e Berger \(2011\)](#), o objetivo de um teste de hipóteses é decidir, baseado em uma amostra da população, qual, entre duas hipóteses complementares, é verdadeira, se a hipótese nula H_0 ou a hipótese alternativa H_1 .

[Casella e Berger \(2011\)](#) define o procedimento de um teste de hipótese como a regra que especifica (i) para quais valores amostrais se decide por aceitar H_0 como verdadeira ou (ii) para quais valores amostrais, H_0 é rejeitada e H_1 é aceita como verdadeira.

Com efeito, a nomenclatura “aceitar H_0 ” é comumente substituída por “não rejeitar H_0 ”. Não rejeitar H_0 é, de fato, distinto de aceitar H_0 , entretanto, como o teste trata de decidir por uma, entre duas afirmações complementares, ao mesmo tempo que se rejeita uma, se aceita obrigatoriamente a outra. Isto quando se considera o resultado do teste como único critério para a tomada de decisão.

O conjunto dos valores amostrais para os quais H_0 será rejeitada é chamado de região de rejeição ou região crítica. Os demais valores formam a região de aceitação, ou de não rejeição.

A abordagem do teste de hipóteses, segundo [Chernick \(2011\)](#), fixa a probabilidade de rejeitar falsamente a hipótese nula e, em seguida, determina um tamanho de amostra que provavelmente resultará na rejeição correta da hipótese nula, quando a diferença

for pelo menos um valor especificado, indicado por δ . Neste sentido, é possível que se decida pela rejeição ou não de H_0 erradamente, assim, define-se os dois tipos de erros de um teste de hipótese:

Erro Tipo 1: Rejeitar H_0 indevidamente, ou seja, afirmar que existe uma diferença δ que na realidade não existe.

A probabilidade de se cometer um erro do tipo 1 é chamada de **nível de significância** (α):

$$\alpha = P(\text{erro tipo 1}) = P(\text{rejeitar } H_0 | H_0 \text{ é verdadeira})$$

Erro Tipo 2: Não rejeitar H_0 indevidamente, ou seja, afirmar que existe uma igualdade quando de fato existe uma diferença.

A probabilidade de se cometer um erro do tipo 2 é denotada por (β):

$$\beta = P(\text{erro tipo 2}) = P(\text{não rejeitar } H_0 | H_0 \text{ é falsa})$$

Considerando que H_0 é falsa, então existe uma diferença δ em relação ao parâmetro em questão, no caso, a média populacional. A probabilidade complementar do erro do tipo 2 ($1 - \beta$) representa a probabilidade de se afirmar que existe uma diferença quando de fato ela existe, sendo o valor ($1 - \beta$) chamado de **poder do teste** de detectar uma diferença que de fato existe. Alternativamente, pode-se definir o poder do teste como sendo a probabilidade de se rejeitar H_0 corretamente ou a probabilidade de aceitar H_0 corretamente.

A tabela 3.1 resume o que foi tratado acima.

Em um teste estatístico espera-se que a probabilidade de se rejeitar H_0 indevidamente seja pequena, ou seja, espera-se que α seja pequeno, este valor é definido antes

Tabela 3.1: Tipos de erro associados à realização dos testes estatísticos e suas respectivas probabilidades

Verdade	Conclusão do teste	
	Não se rejeita H_0	Rejeita-se H_0
H_0 é verdadeira	Decisão correta Probabilidade: $1 - \alpha$	Decisão errada, erro tipo 1 Probabilidade: α
H_0 é falsa	Decisão errada: erro tipo 2 Probabilidade: β	Decisão correta Probabilidade: $1 - \beta$ (poder do teste)

mesmo de se retirar uma amostra de tamanho n da população. Também é esperado que o poder do teste ($1 - \beta$) seja elevado, ou seja, espera-se que o teste tenha alta capacidade de rejeitar H_0 corretamente.

3.3.2 Testes de Hipóteses para a Média Populacional

Os testes de hipóteses para a Média Populacional (THMP) localizam-se na estrutura da Estatística como indica a figura 3.7.

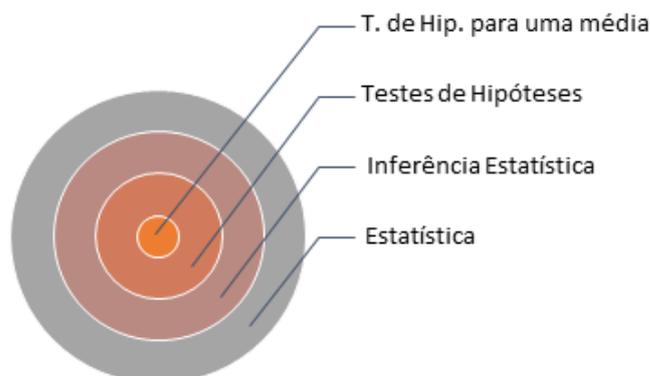


Figura 3.7: Localização dos THMP dentro da estrutura da Estatística.

No caso dos THMP, deseja-se decidir por uma das hipóteses H_0 e H_1 onde, geralmente, a hipótese nula é $H_0 : \mu = \mu_0$ e a alternativa é escolhida entre uma das seguintes possibilidades: $H_1 : \mu \neq \mu_0$; $\mu < \mu_0$ ou $\mu > \mu_0$ onde μ é a média populacional (desconhecida) e μ_0 é um valor de referência para esta média populacional. Aqui, considera-se

que a população seja normalmente distribuída, com média μ e variância σ^2 , denotada por $N(\mu, \sigma^2)$.

Para testar $H_0 : \mu = \mu_0$ contra uma das três hipóteses alternativas, retira-se uma amostra aleatória de tamanho n da população e calcula-se sua média \bar{x} . Caso \bar{x} esteja perto de μ_0 , reforça-se a aceitação da hipótese nula. A proximidade da média amostral \bar{x} de μ_0 é medida em termos de desvios padrão da média amostral $\bar{X} = \frac{\sigma}{\sqrt{n}}$. Quando σ é conhecido, esta medida é chamada de desvio padrão da média. Assim,

$$z = \frac{\bar{x} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

z é chamada estatística do teste.

Estabelece-se o que quer dizer “distante” através do nível de significância α , escrito em termos de Z . Assim, para qualquer \bar{X} escrito em termos de Z , que tenha valor igual ou mais extremo do que Z_α , terá um valor considerado significativamente distante de μ_0 , sugerindo que se rejeite H_0 , ou seja, \bar{X} estará em uma região da distribuição populacional $N(\mu, \sigma^2)$ chamada de **região de rejeição de H_0** .

Costa (2012) chama a região de aceitação (ou não rejeição) da hipótese nula de “intervalo de confiança” e pondera que ele pode ser encarado como um conjunto de hipóteses aceitáveis, onde qualquer hipótese H_0 que esteja fora deste intervalo deve ser rejeitada e as que estiverem dentro do intervalo devem ser aceitas.

Para cada uma das possibilidades de hipótese alternativa, temos as seguintes regiões de rejeição, considerando σ^2 conhecida:

- $\mu < \mu_0$: A região de rejeição será dada por todo \bar{x} que satisfaz

$$Z \leq -Z_\alpha$$

$$\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq -Z_\alpha$$

$$\bar{x} \leq \mu_0 - Z_\alpha \frac{\sigma}{\sqrt{n}}$$

- $\mu > \mu_0$: A região de rejeição será dada por todo \bar{x} que satisfaz

$$Z \geq Z_\alpha$$

$$\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \geq Z_\alpha$$

$$\bar{x} \geq \mu_0 + Z_\alpha \frac{\sigma}{\sqrt{n}}$$

- $\mu \neq \mu_0$: A região de rejeição será dada por todo \bar{x} que satisfaz

$$|Z| \geq Z_{\alpha/2}$$

$$\left| \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \right| \geq Z_{\alpha/2}$$

$$\frac{|\bar{x} - \mu_0|}{\frac{\sigma}{\sqrt{n}}} \geq Z_{\alpha/2}$$

$$|\bar{x} - \mu_0| \geq \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}$$

Faz sentido tomar-se $\alpha/2$ no caso de um teste bilateral ($\mu \neq \mu_0$) quando se pressupõe uma distribuição subjacente simétrica.

Graficamente, as regiões de rejeição apresentam-se como na figura 3.8.

Para o caso de não se conhecer σ^2 , recorre-se ao Teorema Central do Limite que afirma que a razão $\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ tem uma distribuição aproximadamente normal com $N(0,1)$, para um tamanho amostral n suficientemente grande.

Hogg et al. (2013) pondera que quando a distribuição subjacente é unimodal, simétrica e contínua, a aproximação é muito boa mesmo para pequenas amostras. Mas, se tal distribuição tiver acentuada assimetria ou for discreta, um tamanho amostral maior

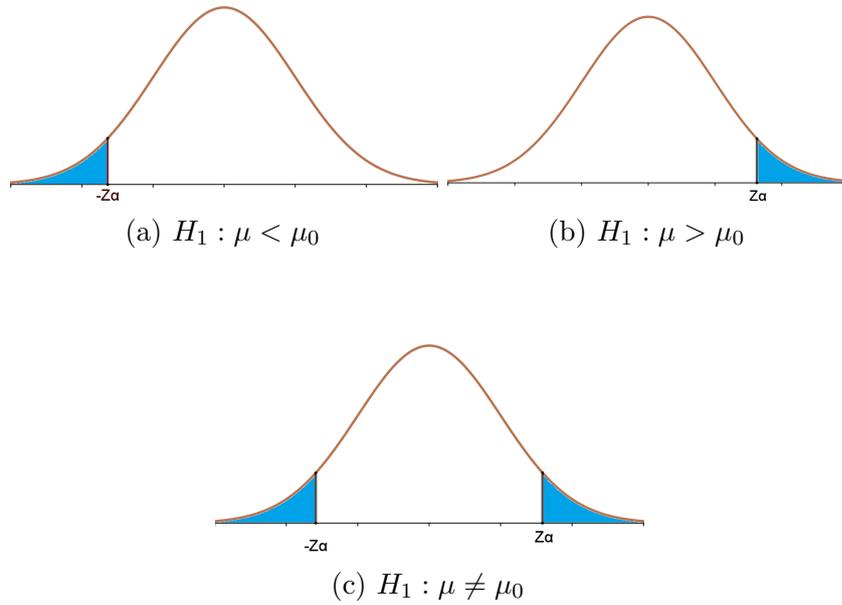


Figura 3.8: Regiões de rejeição de H_0 para cada tipo de H_1 .

é requerido, porém, em ambos os casos, um tamanho de pelo menos $n = 30$ costuma ser adequado.

Assim, se σ^2 for desconhecida e $n \geq 30$, usa-se o fato de $\frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$ ter uma distribuição aproximadamente normal com $N(0,1)$, mesmo que a distribuição subjacente não seja normal. [Hogg et al. \(2013\)](#) observa que se a distribuição subjacente tenha sério problema de assimetria ou que contenha *outliers*, a maioria dos estatísticos prefere uma quantidade amostral maior (50 ou mais) e ainda corre-se o risco de se ter maus resultados.

Não havendo informação sobre a variância ou média da distribuição, mas considerando que tenha uma distribuição aproximadamente normal, faz-se uso do fato de que a razão $T = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$ possui uma distribuição t com $n - 1$ graus de liberdade, onde S^2 é usualmente o estimador não enviesado para σ^2 .

Para cada uma das possibilidades de hipótese alternativa, temos as seguintes regiões de rejeição, considerando o uso da estatística t (variância desconhecida):

- $\mu < \mu_0$: A região de rejeição será dada por todo \bar{x} que satisfaz

$$t \leq -t_\alpha$$

ou

$$\bar{x} \leq \mu_0 - t_\alpha \frac{s}{\sqrt{n}}$$

- $\mu > \mu_0$: A região de rejeição será dada por todo \bar{x} que satisfaz

$$t \geq t_\alpha$$

ou

$$\bar{x} \geq \mu_0 + t_\alpha \frac{s}{\sqrt{n}}$$

- $\mu \neq \mu_0$: A região de rejeição será dada por todo \bar{x} que satisfaz

$$|t| \geq t_{\alpha/2}$$

ou

$$|\bar{x} - \mu_0| \geq t_{\alpha/2} \frac{s}{\sqrt{n}}$$

Se for o caso de não se conhecer a distribuição subjacente, a razão $T = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ será aproximadamente t distribuída com uma boa aproximação, entretanto, se a distribuição tiver grande grau de assimetria ou não for unimodal, [Hogg et al. \(2013\)](#) sugere o uso de um teste não paramétrico.

[Morettin e Bussab \(2017\)](#), apresenta uma outra abordagem para a avaliação da hipótese nula que diz respeito ao cálculo da probabilidade de se obter o valor da estatística, ou valor mais extremo do que o observado, considerando a hipótese nula verdadeira, ou seja, trata-se de calcular a probabilidade de se encontrar valores mais

desfavoráveis a H_0 do que o observado. Chama-se esta probabilidade de *nível descritivo* ou *p-valor*. Assim, quanto menor for o p-valor, maior será a evidência contra H_0 .

Para o caso dos THMP, pode-se definir o p-valor como

$$p = P(Z > |z_0|)$$

ou

$$p = P(T > |t_0|)$$

considerando Z e T as distribuições passíveis de serem escolhidas para o teste e z_0 e t_0 são os valores amostrais de tais distribuições.

O p-valor pode ser apresentado, para cada uma das possibilidades de hipóteses alternativas contra $H_0 : \mu = \mu_0$, como segue:

- $\mu > \mu_0$ e σ^2 **conhecida**: deseja-se saber qual a probabilidade de uma estimativa igual, ou maior, do que \bar{x}_0 ter provido de uma população de média igual a μ_0 .

$$\text{p-valor} = P(\bar{x} \geq \bar{x}_0 | \mu_0) = P\left(Z \geq \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right)$$

- $\mu < \mu_0$ e σ^2 **conhecida**: deseja-se saber qual a probabilidade de uma estimativa igual, ou menor, do que \bar{x}_0 ter provido de uma população de média igual a μ_0 .

$$\text{p-valor} = P(\bar{x} \leq \bar{x}_0 | \mu_0) = P\left(Z \leq \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right)$$

- $\mu \neq \mu_0$ e σ^2 **conhecida**: deseja-se saber qual a probabilidade de uma estimativa diferente de \bar{x}_0 ter provido de uma população de média igual a μ_0 .

$$\text{p-valor} = 2P(\bar{x} \neq \bar{x}_0 | \mu_0) = 2P\left(|Z| \geq \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right)$$

- $\mu > \mu_0$ e σ^2 **desconhecida**, mas $n > 30$: deseja-se saber qual a probabilidade de uma estimativa igual, ou maior, do que \bar{x}_0 ter provido de uma população de média igual a μ_0 .

$$\text{p-valor} = P(\bar{x} \geq \bar{x}_0 | \mu_0) = P\left(Z \geq \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}\right)$$

- $\mu < \mu_0$ e σ^2 **desconhecida**, mas $n > 30$: deseja-se saber qual a probabilidade de uma estimativa igual, ou menor, do que \bar{x}_0 ter provido de uma população de média igual a μ_0 .

$$\text{p-valor} = P(\bar{x} \leq \bar{x}_0 | \mu_0) = P\left(Z \leq \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}\right)$$

- $\mu \neq \mu_0$ e σ^2 **desconhecida**, mas $n > 30$: deseja-se saber qual a probabilidade de uma estimativa diferente de \bar{x}_0 ter provido de uma população de média igual a μ_0 .

$$\text{p-valor} = 2P(\bar{x} \neq \bar{x}_0 | \mu_0) = 2P\left(|Z| \geq \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}\right)$$

- $\mu > \mu_0$ e σ^2 **desconhecida** e $n < 30$: deseja-se saber qual a probabilidade de uma estimativa igual, ou maior, do que \bar{x}_0 ter provido de uma população de média igual a μ_0 .

$$\text{p-valor} = P(\bar{x} \geq \bar{x}_0 | \mu_0) = P\left(T \geq \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}\right)$$

- $\mu < \mu_0$ e σ^2 **desconhecida** e $n < 30$: deseja-se saber qual a probabilidade de uma estimativa igual, ou menor, do que \bar{x}_0 ter provido de uma população de média igual a μ_0 .

$$\text{p-valor} = P(\bar{x} \leq \bar{x}_0 | \mu_0) = P\left(T \leq \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}\right)$$

- $\mu \neq \mu_0$ e σ^2 **desconhecida** $n < 30$: deseja-se saber qual a probabilidade de uma estimativa diferente de \bar{x}_0 ter provido de uma população de média igual a μ_0 .

$$\text{p-valor} = 2P(\bar{x} \neq \bar{x}_0 | \mu_0) = 2P\left(|T| \geq \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}\right)$$

Desta forma, o nível de significância α pode servir como referência na comparação com o p-valor, propondo-se a rejeição de H_0 sempre que o p-valor for menor do que α , tal como ilustrado na figura 3.9.

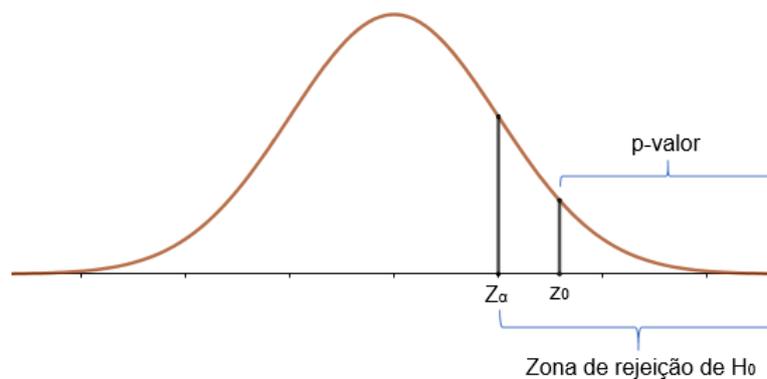


Figura 3.9: Exemplo de p-valor menor do que α .

3.4 O Campo Conceitual dos Testes de Hipóteses na Área da Saúde

Vergnaud (2017) pondera que para analisar o desenvolvimento das competências e das conceitualizações do sujeito nos diferentes registros de sua atividade, é indispensável fragmentar objetos de estudo menores do que a experiência global. Assim, é preciso designar subcampos da experiência em torno das ideias de situação e conceito, ou seja, um campo conceitual (cc). A este respeito, Moreira (2002) lembra que Vergnaud considera o cc como uma unidade de estudo para dar sentido às dificuldades observadas na conceitualização do real.

No âmbito desta pesquisa, há a necessidade de se delinear o CC dos Testes de Hipóteses para Média Populacional (THMP) inserido em um campo mais geral, o dos Testes de Hipóteses (TH). Assim, o delineamento do cc em estudo (cc dos THMP) se estabelece em duas etapas, uma mais geral, descrita nesta seção (3.4) e outra mais específica, dada na próxima seção (3.4.1).

Garfield (1995) já apontava que o ensino de Estatística pode ser mais efetivo se os professores determinarem o que eles realmente querem que os estudantes saibam e façam ao final do seu curso, o que coaduna-se com a ideia de delineamento prévio do CC. Os conteúdos sob os quais se assenta o CC envolvido nesta pesquisa, bem como a ordem típica (mas não única) de construção de tais conteúdos, estão apresentados no mapa conceitual da figura 3.10, e pode ser visto de maneira pormenorizada no apêndice C.

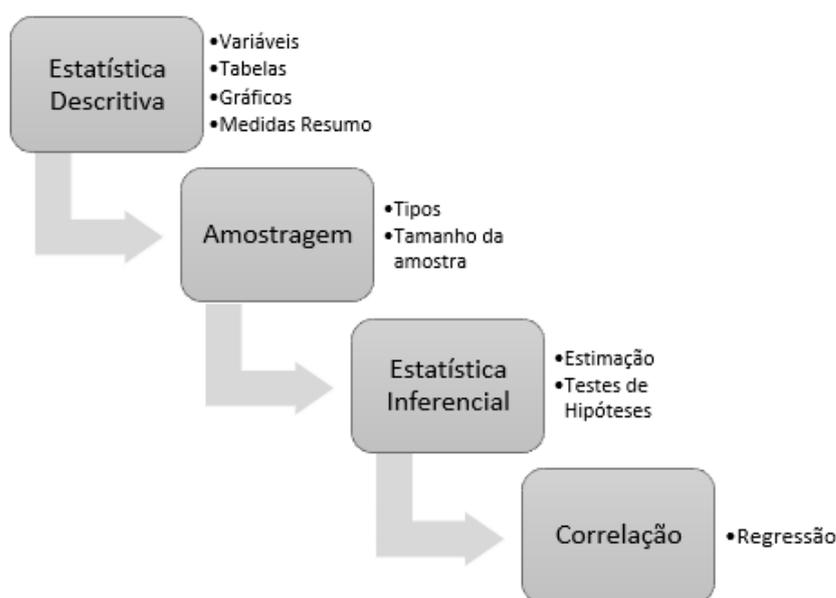


Figura 3.10: Mapa conceitual do campo conceitual que envolvem o Teste de Hipóteses para a Média Populacional (THMP).

O campo conceitual científico do conceito de Teste de Hipóteses (CCTH) é consti-

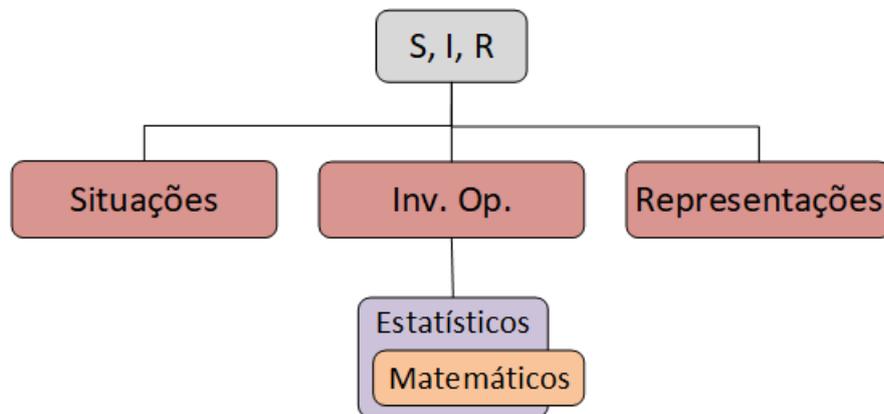


Figura 3.11: Constituintes do CC dos Testes de Hipóteses.

tuído por um conjunto de situações (S), de invariantes operatórios (I) (estatísticos (IE) e matemáticos (IM)) e por um conjunto de representações (R), como mostra a figura 3.11.

O conjunto das situações $S = \{F \cup P\}$ inclui fenômenos (F) e problemas (P) estatísticos que descrevem e dão sentido ao conceito de Teste de Hipóteses relativos à área da saúde como por exemplo, o teste de eficácia de um tratamento. Tratamento é o método, elemento ou material, cujo efeito se deseja medir ou comparar em um experimento [Banzatto e Kronka \(1992\)](#), como um medicamento, a dose de um produto, um procedimento, etc.

O conjunto dos invariantes operatórios $I = \{IE \cup IM\}$, cientificamente aceitos, aplicáveis às situações, dão significado ao conceito de Teste de Hipóteses por meio de suas propriedades, relações e transformações. O conjunto dos invariantes operatórios estatísticos (IE) tem como elementos, as operações e propriedades que se relacionam com os conceitos estatísticos como as medidas estatísticas de tendência central e variabilidade, a probabilidade, o desenho do experimento, a amostragem, o levantamento de dados, a classificação das variáveis, a escolha do teste, etc. que são sujeitos aos princípios estatísticos como o teorema do limite central, a função densidade de probabilidade, os estimadores ótimos, os modelos de probabilidade e de distribuição de frequências. O

conjunto dos invariantes operatórios matemáticos (IM) tem por elementos os significados matemáticos de equação, função, integral, área, etc. e envolvem seus conceitos, operações, propriedades e teoremas principais. É importante salientar que os IM estão relacionados aos IE mas nem sempre dizem respeito aos conceitos estatísticos e sim aos matemáticos, como por exemplo, a resolução de uma equação de primeiro grau que aparece quando se procura o valor do estimador média aritmética no Teste de Hipóteses para a Média Populacional (THMP): a equação em si não depende do conceito estatístico que a gerou, mas dos conceitos matemáticos próprios das equações, ou seja, a equação passa a ter vida própria, essencialmente matemática, e retorna a ter um sentido estatístico quando da interpretação do resultado que ela gerou.

O conjunto das representações simbólicas (R) tem por elementos os significantes estatísticos e matemáticos tanto das situações (S) quanto dos invariantes operatórios (I), de tal forma que as representações permitem externar de maneira coerente e concisa os significados relativos ao campo conceitual de Teste de Hipóteses. Estas representações são de caráter algébrico (quando se escreve uma fórmula ou uma conta qualquer, por exemplo), gráfico (quando se esboça um gráfico, por exemplo), geométrico (quando se hachura a área de uma figura, por exemplo), pictórico (quando se desenha uma figura representativa da relação entre população e amostra, por exemplo) e de escrita natural envolvendo as fórmulas matemáticas, os gráficos, as distribuições de frequências, as hipóteses do teste, a zona crítica em um gráfico de distribuição, etc.

A seguir tem-se a pormenorização do CC dos THMP. Trata-se da exploração prévia dos conceitos constituintes do CC, considerando as dificuldades e erros comuns dos estudantes, alguns Invariantes Operatórios (IO) cientificamente aceitos que são cruciais à teoria, exemplos de situações que dão sentido a estes conceitos, suas notações e representações típicas e alguns de seus desdobramentos.

A seguir apresenta-se um delineamento do conhecimento dos Testes de Hipóteses para a Média Populacional (THMP), considerando-se o seu lugar dentro da estatística

e procurando explicitar os diversos conceitos que o envolvem e são necessários pra o seu entendimento na perspectiva do profissional da saúde em nível de graduação, que está limitado por uma base matemática que normalmente não contempla o Cálculo Diferencial e Integral, necessária para o desenvolvimento mais aprofundado do conteúdo, além de limitações na carga horária da disciplina de Bioestatística nos currículos.

3.4.1 O Campo Conceitual dos Testes de hipóteses para a Média Populacional

Conforme a citação de Vergnaud em [3.1](#), a TCC auxilia o professor no estabelecimento das situações e das intervenções didáticas baseando-se no conteúdo específico. Para tal, o delineamento do Campo Conceitual do Teste de Hipóteses para a Média Populacional, considerando as situações que lhe dão sentido, é determinante para o entendimento dos esquemas alocados pelos Alunos em tais situações durante o processo de conquista deste Campo Conceitual.

Um campo conceitual (CC) leva em conta um conjunto de situações e um conjunto de conceitos, assim, é preciso delinear os sistemas que entrelaçam os conceitos envolvidos no CC, mesmo reconhecendo a impossibilidade de fazê-lo de maneira completa.

[Moreira \(2002\)](#) ressalta que Vergnaud considera útil falar em distintos CC se eles puderem ser consistentemente descritos, mesmo ele crendo que é praticamente impossível estudar as coisas separadamente, é por isto mesmo que se deve fazer recortes nos quais os campos conceituais são unidades de estudo frutíferas para dar sentido aos problemas de aquisição e às observações feitas em relação à Conceitualização, além da necessidade de se dar toda atenção aos aspectos conceituais dos esquemas de pensamento e a análise conceitual das situações nas quais os aprendizes desenvolvem seus esquemas de pensamento.

Vergnaud normalmente apresenta (e.g. [Vergnaud \(2003b, 2008\)](#)) um campo conceitual considerando os conceitos em si, as situações que os envolvem, as possibilidades

para além do foco que se está trabalhando, as maneiras sob as quais os alunos costumam abordar os problemas e como o próprio Vergnaud supõe que seriam e as dificuldades que os alunos têm em cada parte das situações.

A etapa descrita a seguir ajudou na composição dos instrumentos de coleta de dados desta Pesquisa (Cap. 4), pois foi possível identificar possíveis invariantes operatórios necessários ao processo do THMP, resumidos nas tabelas 3.2 e 3.3, além daqueles típicos, porém errôneos, que os estudantes podem vir a apresentar. Este exercício pedagógico fundamental foi baseado em artigos (e.g. [Teresita Terány e Jesica Ciminari \(2019\)](#); [Batanero et al. \(2012\)](#)) que remetem às dificuldades gerais dos alunos bem como na experiência do professor pesquisador, que traz as questões didáticas locais mais comuns que podem impactar na pesquisa. Tais questões foram estabelecidas antes das aplicações das atividades da pesquisa.

Neste sentido, [Teresita Terány e Jesica Ciminari \(2019\)](#) pesquisaram 275 avaliações finais de bioestatística referentes também a teste de hipóteses e encontraram que os erros mais frequentes referem-se ao estabelecimento das hipóteses, à diferenciação entre parâmetro e estatística, à regra de decisão adotada e à interpretação incorreta do teste em termos do problema proposto. Já o trabalho de [Batanero et al. \(2012\)](#) traz que a maioria das dificuldades de alunos universitários está na compreensão do nível de significância.

Com isto em mente, passa-se ao delineamento do CC dos THMP.

O CC dos THMP consiste de todas as situações que podem ser analisadas como problemas onde é preciso interpretar os efeitos de rejeitar ou não uma afirmação sobre a média populacional (H_0) tomando-se por referência os dados observados em uma amostra, a estatística do teste (Z ou t) e um critério de decisão (comparação com α).

Na sequência apresenta-se uma situação típica do CC dos THMP e o seu delineamento:

A Sociedade Brasileira de Hipertensão (SBH) preconiza⁷ que “a pressão arterial ideal para a minimização do risco de problemas cardiovasculares situa-se abaixo de 120/80 mmHg”. Uma amostra aleatória contendo 52 estudantes de uma universidade, apontou uma média amostral para a pressão arterial sistólica (PAS) de 123 mmHg e um desvio padrão de 8 mmHg. Considerando que a PAS possui uma distribuição populacional normal, há evidência de que a PAS de todos os estudantes desta instituição seja diferente de 120 mmHg, a um nível de significância de 5%?

Para dar conta desta situação o estudante deve possuir um entendimento mínimo do processo de inferência que se quer conduzir (Figura 3.12). Aqui, as ideias de População e Amostra, de Parâmetro e Estatística precisam estar bem claras, sob pena de não se compreender o todo do processo de inferência. Contudo, este entendimento fica naturalmente implícito, a menos que se peça ao estudante que o exponha de alguma maneira, visto que sequer existe uma estrutura simbólica típica deste passo.

Além destas ideias gerais, é necessário que o estudante tenha presente que o objetivo do THMP é, especificado um valor μ_0 , avaliar a relação dos dados com esse valor, ou seja, se está supondo que $H_0 : \mu = \mu_0$ é verdadeira e que H_0 se refere a um modelo probabilístico, uma idealização (Cox, 2006, p. 31).

Este objetivo pode ser colocado de maneira procedimental, de mais de uma forma, dentre elas:

- Considerando H_0 verdadeira, ou seja, que a média da população é igual ao valor de referência, é razoável em termos probabilísticos, obter-se uma média amostral igual a \bar{x} em uma amostra aleatória de n indivíduos desta população? (Adaptado de (Callegari-Jacques, 2003, p. 56))
- Supondo que a amostra provenha de uma população cuja média é μ_0 , qual a probabilidade de se obter a média amostral \bar{x} (ou outra mais extrema) para uma

⁷<http://www.sbh.org.br/informacoes.html#o-que-e>

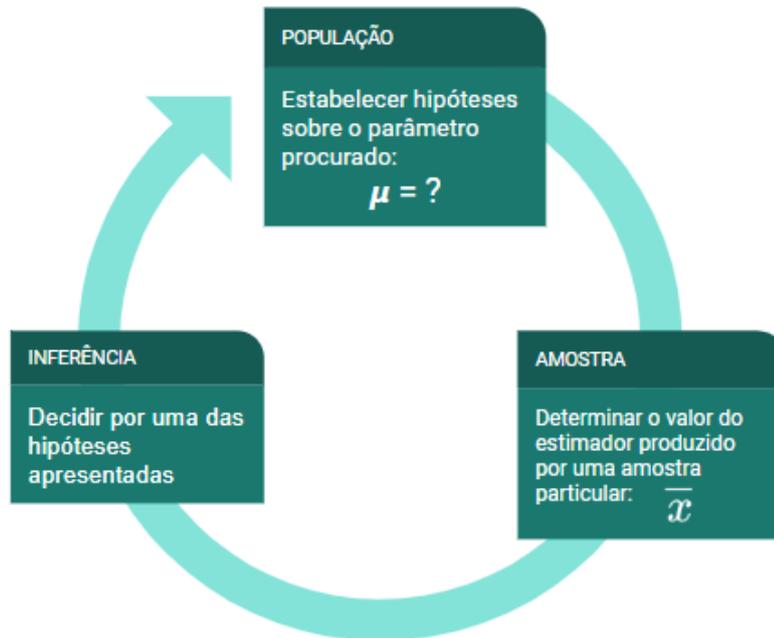


Figura 3.12: Ideia mínima esperada que o estudante tenha sobre o processo do THMP.

amostra de tamanho n ?

- Qual a probabilidade de se obter o que eu obtive se não há relação entre as variáveis (Dancey e Reidy, 2018, p. 162).

Neste ponto inicial da resolução da situação é comum estabelecer-se a hipótese nula e alternativa por um processo puramente mecânico, já que as hipóteses do teste contam com símbolos típicos que os alunos apressam-se em utilizá-los. Por outro lado, é também comum observar a não explicitação das hipóteses nula e alternativa no início do processo, retomando esta importante etapa somente no final do teste.

Normalmente, a afirmação a ser testada indica que a amostra observada provém de uma população cuja média é um determinado valor de referência, neste exemplo, tem-se:

$$H_0 : \mu = 120$$

$$H_1 : \mu \neq 120$$

O ato de estabelecer a hipótese nula pode revelar dificuldades de interpretação e do processo em si, como quando o estudante escreve $H_0 : \bar{x} = 120$, onde parece não perceber que as hipóteses dizem respeito ao parâmetro e não à amostra, ou quando ele sequer estabelece a hipótese.

A hipótese nula aparece no início da resolução de problemas envolvendo THMP e são referenciadas ao final da resolução, sendo ela a hipótese a ser testada, configurando seu papel crucial para a resolução e interpretação deste tipo de problema, entretanto, acontece dos estudantes não iniciem a resolução com o estabelecimento da hipótese nula, mesmo que a tomada de decisão refira-se a ela, de certa forma, invertendo o processo de resolução e fazendo supor que, ao fazê-lo, o estudante não tem em mente a hipótese nula durante o processo em detrimento de uma resolução mais procedimental e menos interpretativa.

A escolha da hipótese alternativa pode ser um processo delicado para o estudante visto que ela depende da interpretação mais detalhada da situação. O estudante corre o risco de escolher $H_1 : \mu > 120$ justificando que a média amostral foi 123 mmHg, ou escolher $H_1 : \mu < 120$ por julgar que a SBH preconiza este valor como o mais saudável. Entretanto, a hipótese nula correta considera que a população tem uma PAS ideal, restando para H_1 o antagônico deste fato, ou seja, $H_1 : \mu \neq 120$, dado que não há indicação de que a PAS seja maior ou menor do que o valor de referência.

No processo do TH, o estudante precisa estar ciente de que a hipótese nula não será aceita ou refutada de forma categórica, senão será estabelecida uma probabilidade de ocorrência dos valores amostrais, considerando *a priori* que a hipótese nula é verdadeira e, em seguida, comparando-a com o nível de significância α , assim, a expressão “verdadeira” ou “falsa” é inapropriada para H_0 e a expressão típica “rejeitar” ou “não rejeitar” a hipótese nula deve ser entendida como provável ou não que a amostra observada provenha da população com a média em questão.

Neste ponto, o CC dos THMP entrelaça-se com os sistemas de conceitos de distri-

buição de frequência e de probabilidade, probabilidade condicional e amostragem, além do conceito de erro do tipo 1, componente do conceito de nível de significância que, a esta altura, deve estar bem compreendido.

Os conceitos subjacentes aos TH de, variável aleatória e função densidade de probabilidade (fdp) são cruciais neste ponto, ao menos em nível operatório, levando-se em conta que estes conceitos não constam na ementa da disciplina de bioestatística dos cursos nos quais esta pesquisa toma suas amostras. A este respeito, espera-se que os estudantes compreendam o conceito de variável aleatória como sendo o conjunto dos resultados numéricos de um evento que não se pode conhecer um resultado particular. Por exemplo, ao se jogar um dado, se conhece as possibilidades de resultados mas não se sabe qual será o resultado particular de um lançamento individual. Simbolicamente, utiliza-se uma letra maiúscula para designar uma variável aleatória. Para o conceito de função densidade de probabilidade, espera-se que o estudante compreenda-o como o valor da probabilidade de ocorrência de um intervalo de elementos possíveis, representado por $P(a \leq X \leq b) = c$, ou seja, c é a probabilidade de que a variável aleatória assuma valores entre a e b , com $X, a, b \in \mathfrak{R}$. Por consequência, a probabilidade condicional representada por $P(E_2|E_1)$, deve ser entendida como a probabilidade de ocorrência do evento 2 (E_2) depois que o evento 1 (E_1) ocorreu.

Chega o momento da escolha da estatística a ser utilizada. Para o exemplo em questão, dado que σ é desconhecido mas $n > 30$, usa-se a estatística Z . A figura 3.13 mostra uma representação da região de rejeição de H_0 . Há muito de implícito nesta etapa, entretanto, não é raro que os alunos indiquem, com setas apontando de $n > 30$ para Z , por exemplo, o raciocínio que os levou a decidir sobre a estatística correta. Dificilmente o aluno relaciona explicitamente esta etapa com a escolha de um modelo teórico de distribuição.

O procedimento continua com o cálculo da estatística do teste, para o exemplo,

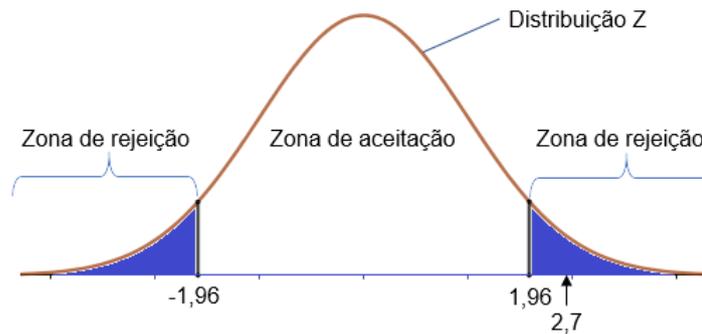


Figura 3.13: Zona de rejeição para uma distribuição Z e $\alpha = 5\%$.

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \Rightarrow z = \frac{123 - 120}{\frac{8}{\sqrt{52}}} \Rightarrow z = 2,7$$

E, como $z > 1,96$, rejeita-se a hipótese nula. Aqui o entendimento pode dar-se num sentido puramente algébrico e mecânico, com o estudante memorizando a fórmula e desenvolvendo-a para a determinação do valor de z , porém, é preciso que ele compreenda que se está calculando a quantidade de erros padrão que a média amostral se distancia da média de referência, explicitando assim, a ideia básica do THMP.

Caso opte-se pela abordagem do p-valor, o qual pode indicar o quão incompatível com um modelo de probabilidade específico são os dados amostrais, usa-se a comparação do p-valor com o nível de significância. Se o p-valor for menor do que α , significa que a estatística do teste está na região de rejeição de H_0 . Simbolicamente, para o exemplo em questão:

$$p - \text{valor} = P(Z > |z|) + P(Z < -|z|) = 0,0069$$

E, como $p - \text{valor} < \alpha$, rejeita-se a hipótese nula.

Ao final do teste, mais uma vez o processo mecânico pode fazer com que um conhecimento superficial confunda-se com um conhecimento mais profundo, visto que a alocação simples da estatística do teste na zona de rejeição da hipótese nula é trivial,

porém, o seu entendimento não o é. O mesmo acontece ao utilizar-se a abordagem do p-valor. Aqui, ressurge a importância do estabelecimento de momentos didáticos onde há a explicitação, pelo aluno, dos conceitos envolvidos, reforçando que o saber fazer pode estar distante do saber dizer.

Em um curso introdutório de estatística para a área da saúde é comum distinguir duas abordagens para os THMP, a da região crítica e a do p-valor, mesmo que a segunda decorra da primeira. Abaixo, seguem as etapas de desenvolvimento das duas abordagens e a ação esperada ao desenvolvê-las:

Tabela 3.2: Campo Conceitual dos Teste de Hipóteses: método da região crítica

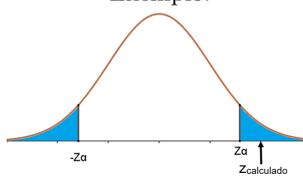
	Etapa	Ação esperada na etapa	Representação usual
1	Leitura e interpretação do item	Identificar o tipo de teste: bilateral ou unilateral.	Texto
2	Determinação das hipóteses	Enunciar as hipóteses nula e alternativa	$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$ ou $H_1 : \mu > \mu_0$ ou $H_1 : \mu \neq \mu_0$
3	Determinação do nível de significância	Enunciar o nível de significância α desejado	Usualmente: $\alpha = 1\%, 5\%$ ou 10%
4	Escolha da estatística do teste	Escolher entre Z ou T	Se σ^2 conhecido (ou desconhecido mas $n > 30$) usa-se Z. Se σ^2 desconhecido e $n \leq 30$ usa-se T
5	Delimitação da zona de rejeição (e por consequência a de não rejeição) de H_0	A partir da hipótese alternativa e da distribuição escolhidas, utilizar as tabelas de Z ou T e o nível de significância estabelecido para delimitar a zona de rejeição de H_0	$z_{\alpha \text{ tabelado}}$ ou $t_{\alpha, (n-1), \text{tabelado}}$
6	Cálculo do valor da estatística do teste	Considerando as escolhas feitas até aqui e dos dados amostrais, determinar a estatística do teste	$(z \text{ ou } t)_{\text{calculado}} = \frac{\bar{x} - \mu_0}{\frac{(\sigma \text{ ou } s)}{\sqrt{n}}}$
7	Verificação da localização da estatística do teste	Verificar se o valor da estatística do teste pertence ou não à zona de rejeição de H_0	<p>Exemplo:</p>  <p>ou $z_{\text{calculado}} > Z_{\text{tabelado}}$</p>
8	Tomada de decisão	Rejeitar ou não rejeitar H_0 baseando-se na verificação se a estatística do teste pertence ou não à zona de rejeição de H_0	<p>Exemplo: se $z_{\text{calculado}} > Z_{\text{tabelado}}$ então rejeita-se H_0 e aceita-se H_1, ou seja, a média amostral é significativamente maior do que o parâmetro de referência</p>

Tabela 3.3: Campo Conceitual dos Teste de Hipóteses: método do p-valor

	Etapa	Ação esperada na etapa	Representação usual
1	Leitura e interpretação do item	Identificar o tipo de teste: bilateral ou unilateral.	Texto
2	Determinação das hipóteses	Enunciar as hipóteses nula e alternativa	$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$ ou $H_1 : \mu > \mu_0$ ou $H_1 : \mu \neq \mu_0$
3	Determinação do nível de significância	Enunciar o nível de significância α desejado	Usualmente: $\alpha = 1\%, 5\%$ ou 10%
4	Escolha da estatística do teste	Escolher entre Z ou T	Se σ^2 conhecido (ou desconhecido mas $n > 30$) usa-se Z. Se σ^2 desconhecido e $n \leq 30$ usa-se T
5	Calculo do p-valor	A partir da hipótese alternativa e da distribuição escolhidas, calcular o p-valor	Desenvolvimento de alguma das opções para o p-valor da página 76
6	Comparação do p-valor com α	Verificar se o p-valor é maior ou menor do que α	$p\text{-valor} < \alpha$ ou $p\text{-valor} \geq \alpha$.
7	Tomada de decisão	Rejeitar ou não rejeitar H_0 a partir da comparação do p-valor com o α	Se $p\text{-valor} < \alpha$ rejeita H_0 , caso contrário, não rejeita H_0 .

O p-valor é largamente utilizado na pesquisa da área da saúde, em detrimento do caminho da região crítica, provavelmente pela simplicidade na tomada de decisão. Além disto, com a popularização dos computadores e dos *softwares* estatísticos, o cálculo do p-valor (que nem sempre é fácil) está ao alcance dos pesquisadores de todas as áreas, mesmo os sem habilidades com o cálculo de probabilidades. Somando-se a isto, o uso do caminho da região crítica necessita da verificação do pertencimento da estatística

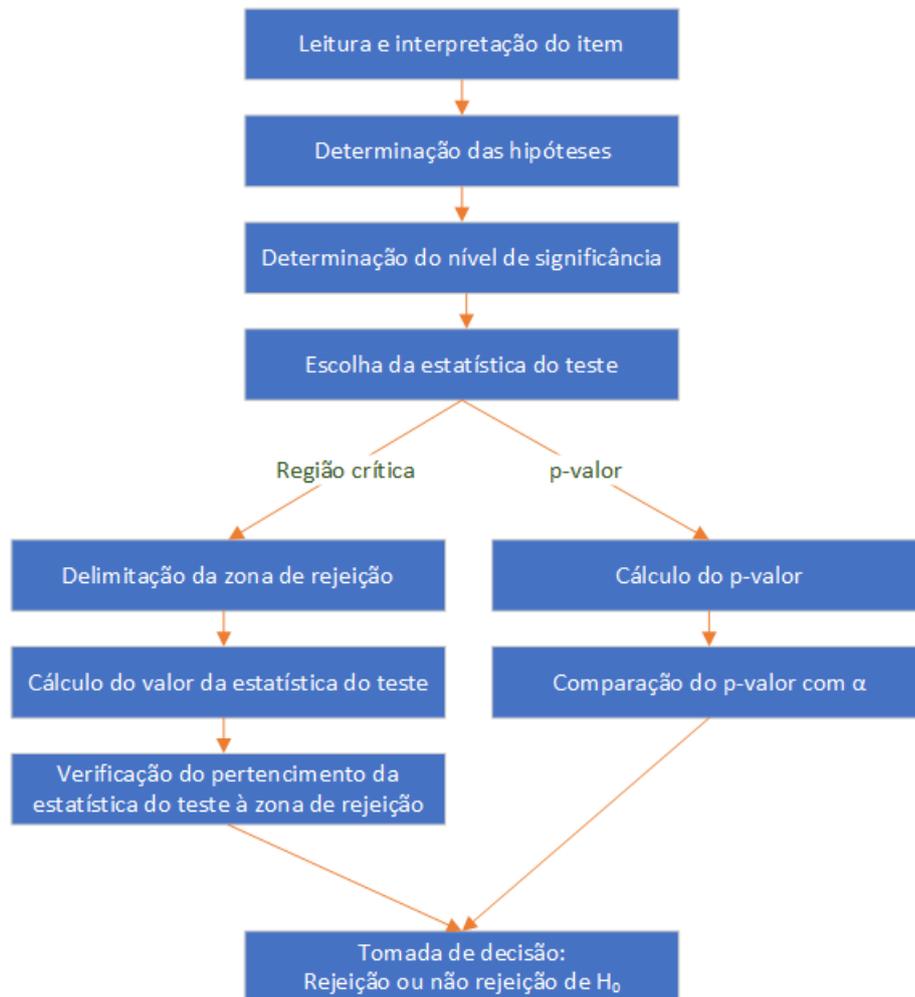


Figura 3.14: Opções de caminhos a seguir para o desenvolvimento dos Testes de Hipóteses.

do teste (também calculada pelos *softwares* estatísticos) à zona de rejeição, a qual não fica evidente nas saídas de muitos *softwares* estatísticos, como se pode ver nas figuras 3.15 e 3.16.

- 1 -	
Tamanho da amostra	58
Variância	50.5653
Desvio Padrão da população	7.1109
Erro padrão da população	0.9337
Média amostral	22.4310
Média da população	23.0000
(Z)=	-0.61
(p) unilateral	0.2709
(p) bilateral	0.5419
Poder (0.05)	0.1491
Poder (0.01)	0.0274
IC 95%(média amostral)=	20.6010 a 24.2611
IC 99%(média amostral)=	20.0221 a 24.8400

Figura 3.15: Exemplo de saída do *software* estatístico Bioestat.

Teste de uma amostra						
	Valor de Teste = 23					
	t	df	Sig. (2 extremidades)	Diferença média	95% Intervalo de Confiança da Diferença	
					Inferior	Superior
EscoreESV	-.609	57	.545	-.569	-2.44	1.30

Figura 3.16: Exemplo de saída do *software* estatístico Bioestat.

Para que o aluno interprete corretamente esta saída (figura 3.15), precisará primeiro definir a região crítica (por exemplo: rejeita-se H_0 se $|z_{calculado}| \geq 1,96$, considerando $\alpha = 0,05$) e em seguida verificar se o valor da estatística do teste dada pelo *software* pertence ou não à região crítica. Neste caso $-0,61$ não pertence à região crítica, logo, não se rejeita a hipótese nula. Outra representação é dada pela figura 3.17.

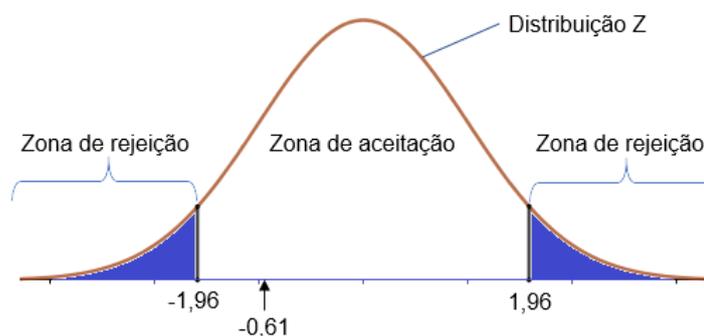


Figura 3.17: Representação da localização da estatística do teste frente a região crítica.

O delineamento do CC dos THMP feito neste apartado, trouxe muito das questões envolvidas no raciocínio, letramento e pensamento estatístico, que são pormenorizados na próxima seção.

3.5 *Problem Based Learning - PBL*

Diante de uma sala de aula pouco atrativa ao estudante, posições que contrapõem o modelo tradicional são necessárias. Este cenário, embora atual, não é novo, como se depreende de [Ribeiro \(2008\)](#).

O PBL originou-se na Escola de Medicina da Universidade Mc Master (Canadá) no final dos anos 1960, inspirado no método de casos de ensino da escola de Direito da Universidade de Harvard (EUA) na década de 1920 e no modelo desenvolvido na Universidade Case Western Reserve (EUA) para o ensino de medicina no anos 1950. A implantação do PBL no contexto educacional original veio em resposta à insatisfação e ao tédio dos alunos frente ao grande volume de conhecimentos percebidos como irrelevantes à prática médica. Esta iniciativa também foi decorrente do fato de seus formandos estarem deixando o curso com muitos conceitos, mas com poucas estratégias e poucos comportamentos associados à aplicação de informações a um diagnóstico. ([Ribeiro, 2008](#), p. 14)

Segundo [Savin-Baden et al. \(2004\)](#), o PBL iniciou depois de anos de pesquisa na busca de melhorias no ensino de medicina, precário na época, desenvolvida por Donald

Woods, creditado por cunhar o termo PBL e por trazer à tona a metodologia criada pelos educadores Barrows e Tamblyn.

Percebe-se que o PBL tem uma preocupação com a motivação do estudante e com questões atitudinais que vão além do conteúdo e avançam ao mundo do trabalho.

Da mesma forma, [Vergnaud \(2003b\)](#) também se preocupa com a formação profissional.

O desenvolvimento da pesquisa em didática, nestes últimos vinte anos, corresponde à conjunção de várias razões. A primeira é a necessidade de formar, num nível de qualificação mais ou menos elevado, uma proporção cada vez maior de indivíduos. ([Vergnaud, 2003b](#), p. 74)

Também é necessário lembrar que, para [Vergnaud \(2003b\)](#), “a parte essencial dos processos cognitivos é a conceitualização”.

Assim, no âmbito desta Pesquisa, o foco reside na formação estatística adequada do profissional da área da saúde, de tal forma que a atitude investigativa própria de sua profissão seja exercitada também por meio do fazer estatístico, durante o processo de conceitualização do conteúdo dos testes de hipóteses para a média populacional (THMP).

A metodologia do PBL pode suprir esta necessidade, dadas suas características, discutidas a seguir.

Segundo [Boud e Feletti \(1997\)](#), a PBL é uma técnica de ensino usada em muitas escolas médicas para facilitar o aprendizado de conceitos científicos no contexto dos casos clínicos, mas não deve ser vista como uma maneira particular de aprendizagem, mas como um aprendizado que possui várias formas diferentes, diferindo de acordo com a natureza da disciplina e os objetivos específicos do programa.

A flexibilidade do PBL permite sua utilização em ambientes educacionais que não têm esta metodologia em seu currículo. No caso desta Pesquisa, o PBL foi aplicado durante um período limitado da disciplina de Bioestatística, como um intervalo de

exceção dentro de um currículo convencional, o que está adequado à ideia de [Ribeiro \(2008\)](#).

[...] o PBL pode ser implantado em todo o currículo (como foi originalmente concebido), em um de dois segmentos paralelos de um mesmo currículo (formato híbrido) ou em uma ou mais disciplinas em um currículo convencional (formato parcial). Há também relatos da utilização pontual do PBL, conhecida como *post-holing*, em determinados momentos de disciplinas baseadas em aulas expositivas, quanto se deseja integrar conhecimentos ou aprofundar determinados tópicos. ([Ribeiro, 2008](#), p. 22)

No PBL, segundo [Duch et al. \(2001\)](#), problemas complexos e reais são usados para motivar os alunos a identificar e pesquisar os conceitos e princípios que precisam saber para trabalhar com esses problemas. Os alunos trabalham em pequenas equipes de aprendizagem, reunindo habilidade coletiva na aquisição, comunicação e integração de informações.

O ingrediente chave do PBL, para [Hadgraft e Holecek \(1995, p. 8\)](#), são:

- a aprendizagem ativa, por meio da criação de perguntas e busca por respostas;
- aprendizagem integrada, pelo estabelecimento de problemas cuja solução exige o conhecimento de várias subáreas;
- aprendizagem cumulativa, por meio de problemas com grau de complexidade crescente até alcançar aqueles geralmente enfrentados por profissionais iniciantes;
- aprendizagem para a compreensão, em vez de para a retenção de informações, pelo provimento de tempo para a reflexão, *feedback* frequente e oportunidades para praticar o que aprendeu.

[Boud e Feletti \(1997\)](#) apresentam a sua ideia do processo de desenvolvimento da PBL da seguinte forma:

- Os alunos recebem um problema (caso, documento de pesquisa, vídeo, ...). Trabalham em grupos para organizarem suas ideias e conhecimentos prévios relacionados com ele.
- Durante a discussão, os alunos definem as “questões de aprendizagem” que delimitam aspectos do problema que eles não compreendem. Esses problemas são registrados pelo grupo e ajudam gerar e focar a discussão. Os alunos são continuamente encorajados a definirem o que eles sabem e, o mais importante, o que eles não sabem.
- Os alunos classificam, por ordem de importância, as questões de aprendizagem geradas. Eles decidem quais questões serão abordadas pelo grupo e quais serão atribuídas a indivíduos que mais tarde ensinarão o resto do grupo. Estudantes e tutor(es) discutem quais recursos serão necessários e onde poderiam ser encontrados.
- Quando os alunos se reagrupam, exploram os problemas de aprendizagem anteriores, integrando seus novos conhecimentos no contexto do problema. Os alunos são encorajados a resumir seus conhecimentos e conectar novos conceitos aos antigos. Eles continuam a definir novas questões de aprendizagem à medida que progredem pelo problema. Eles perceberão que aprender é um processo contínuo e que sempre haverá (mesmo para o professor) questões de aprendizagem a serem exploradas.

As etapas do PBL podem ser resumidas como mostra a figura [3.18](#).

Assim definida, a PBL ajusta-se ao conceito de metodologia ativa, na perspectiva dada por [Valente et al. \(2017\)](#).

A maior parte da literatura brasileira trata as metodologias ativas como estratégias pedagógicas que colocam o foco do processo de ensino e aprendizagem no aprendiz, contrastando com a abordagem pedagógica do ensino

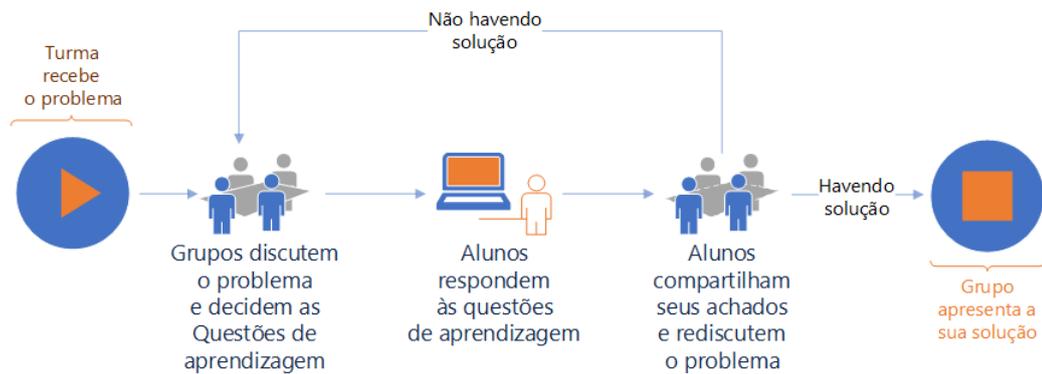


Figura 3.18: Etapas do *Problem Based Learning*.

tradicional, centrada no professor, que transmite informação aos alunos. O fato de elas serem caracterizadas como ativas está relacionado com a aplicação de práticas pedagógicas para envolver os alunos, engajá-los em atividades práticas, nas quais eles são protagonistas da sua aprendizagem. (Valente et al., 2017, p. 463)

Pelo exposto, percebe-se o alinhamento do PBL com os seguintes preceitos de Garfield e Ben-Zvi (2007), vistos na seção 2.3:

- Estudantes aprendem pela construção do conhecimento;
- Estudantes aprendem pelo envolvimento ativo em atividades de aprendizagem;
- Estudantes aprendem a fazer bem somente o que praticam;
- A aprendizagem melhora se os estudantes tomam consciência e confrontam seus erros de raciocínio;
- Estudantes aprendem melhor se recebem *feedbacks* consistentes que auxiliem em sua performance.

Neste ponto, fica exposto que a escolha do problema é fundamental para o PBL. Para os fins desta pesquisa, os problemas escolhidos vieram do campo de estudo dos Alunos e foram construídos a partir de ideias de profissionais destes campos. As turmas

do curso de Psicologia trabalharam com a Escala de Satisfação de Vida (ESV) enquanto que a turma do curso de Nutrição debruçou-se sobre dados referentes à pacientes pós-bariátricos. Ambos os problemas podem ser vistos nos anexos [F](#) e [G](#).

Com estes problemas, procurou-se estar na zona de desenvolvimento proximal (ZDP) de, pelo menos, a maioria dos alunos. [Vergnaud \(2003b\)](#) refere-se da seguinte forma à ZDP.

Vygotski procurou, antes dos outros, definir a margem na qual o educador pode representar seu papel. A “zona proximal de desenvolvimento” é, segundo sua definição, o que a criança é capaz de fazer com a ajuda dos outros e que não poderia fazer sozinha. É essa ideia, talvez, que o leva a distinguir (e, quem sabe, a opor mais do que seria necessário) os aprendizados que resultam principalmente da experiência da criança em seu meio ambiente cotidiano dos que resultam de uma ação intencional do adulto e da instituição escolar. ([Vergnaud, 2003b](#), p. 68)

Parece ser útil que o tutor do PBL tenha em mente a existência de uma região cognitiva no estudante (a ZDP) onde o aprendizado está prestes a acontecer, por sua ação ou pela ação dos outros Alunos. Sugerir situações propensas a estarem nesta região, têm mais chances de serem eficientes no desenvolvimento dos conceitos a que dão sentido. Atenta-se, assim, à importância do conteúdo inserido na ZDP para a TCC:

A teoria dos campos conceituais é uma teoria do desenvolvimento que dá um conteúdo concreto à zona do desenvolvimento proximal, conteúdo conceitual por conteúdo conceitual, tanto é verdade que o problema da apropriação dos conhecimentos é muito dependente deste conteúdo. ([Vergnaud, 2008](#), p. 31)

Para além da proposição adequada das situações que dão sentido ao conceito, ao Tutor cabe a mediação do processo:

[...] atos de mediação do professor vão ter que acompanhar, pelo menos de vez em quando, o trabalho do aluno, jamais totalmente autônomo, seja o que for aquilo que desejamos. Esses atos de mediação tomam muitas vezes a

forma de palavras, mas não somente; um gesto do dedo apontado para o erro pode significar tanto quanto uma explicação. Contudo, as explicações são indispensáveis, e a forma predicativa do conhecimento vem então socorrer a forma operatória que é o “esquema de pensamento”, com uma função vizinha, guardadas as proporções, daquela da linguagem da criança que aprende a falar, e cujas primeiras palavras e enunciados vêm socorrer os primeiros invariantes operatórios construídos por ela, objetos e relações. (Vergnaud, 2008, p. 31)

Desta forma, Vergnaud insere diretrizes importantes no fazer pedagógico do Tutor do PBL ao mesmo tempo que limita a pretensão de um aprendizado totalmente autônomo. Neste sentido, pondera-se que o papel do professor, por ventura Tutor no PBL, não é somente o de organizar e orientar o processo, mas passa tanto pela explicação do conteúdo quanto pela sua postura perante o Educando. Assim, não é demais lembrar, que o fazer pedagógico é função do profissional da educação habilitado para tal, o professor, e não pode ser confundido com um fazer técnico ou administrativo, ou seja, o PBL apoiado pela TCC tem que ser desenvolvido por professor.

Capítulo 4

Metodologia de Pesquisa

A metodologia de pesquisa seguiu uma estratégia integradora entre os paradigmas qualitativo e quantitativo por meio da triangulação metodológica, visando o melhor entendimento do nível de domínio do campo conceitual (CC) dos testes de hipóteses para a média populacional (THMP) de alunos da área da saúde.

As abordagens qualitativa e quantitativa apoiaram-se na análise de dados amostrais levantados por três recursos:

1. O diário de classe permitiu a análise interpretativa descritiva da observação participativa conduzida durante a pesquisa qualitativa desenvolvida para a aplicação do PBL e está melhor descrito na subseção [4.2.1](#).
2. A aplicação de dois itens (questões) abertos ao final da metodologia do PBL permitiu tanto uma análise qualitativa quanto quantitativa para o estabelecimento do nível de apropriação do CC dos THMP. Este instrumento está melhor descrito na subseção [4.2.3](#).
3. A aplicação de um instrumento composto de itens fechados (questões de múltipla escolha), após a aplicação do recurso anterior, permitiu uma análise quantitativa do rendimento dos Estudantes num cenário avaliativo tradicional (prova). Este

instrumento possibilitou a comparação de resultados entre os alunos que desenvolveram a PBL com os que não a desenvolveram. A descrição deste instrumento pode ser vista em 4.2.2.

A figura 4.1 ilustra a metodologia da pesquisa.

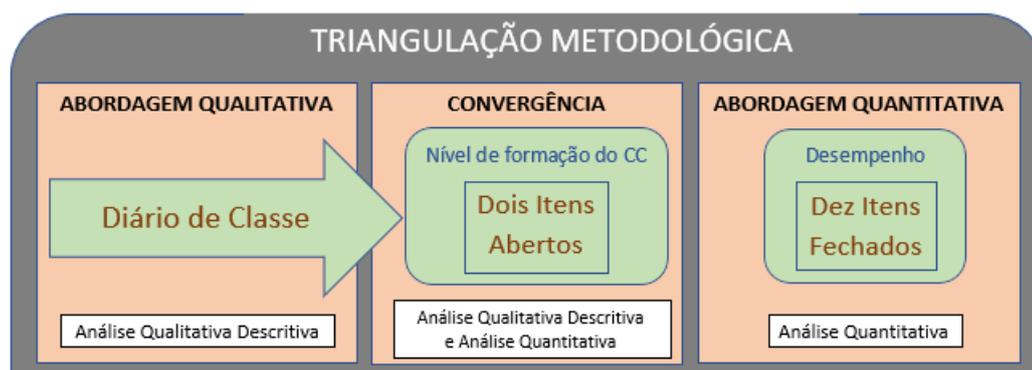


Figura 4.1: Desenho da metodologia de pesquisa.

A partir deste ponto, passa-se a discorrer sobre os elementos principais da metodologia descrita acima. Para [Moreira \(2011\)](#), o interesse central da pesquisa qualitativa está na interpretação dos significados atribuídos pelos sujeitos à suas ações em uma realidade socialmente construída, através da observação participativa, o que neste caso, tomará registro no diário de classe.

Desta forma, a convivência dos sujeitos envolvidos na Pesquisa, em seus distintos papéis de professor e alunos, gera o rico cenário de ensino que pode ser captado pelo diário de classe.

O diário de classe também está adequado à visão de [Schmidt Godoy \(1995\)](#) por preservar o caráter subjetivo de uma experiência didática.

De maneira diversa [à pesquisa quantitativa], a pesquisa qualitativa não procura enumerar e/ ou medir os eventos estudados, nem emprega instrumental estatístico na análise dos dados. Parte de questões ou focos de interesses amplos, que vão se definindo à medida que o estudo se desenvolve. Envolve a obtenção de dados descritivos sobre pessoas, lugares e processos interativos

pelo contato direto do pesquisador com a situação estudada, procurando compreender os fenômenos segundo a perspectiva dos sujeitos, ou seja, dos participantes da situação em estudo. (Schmidt Godoy, 1995, p. 58)

A pesquisa quantitativa, segundo Moreira (2011), procura estudar os fenômenos de interesse da pesquisa em educação geralmente através de estudos experimentais ou correlacionais, caracterizados, primordialmente, por medições objetivas e análises quantitativas.

O instrumento fechado foi aplicado a todos os estudantes, participantes ou não do PBL. Neste sentido, o estudo pode ser caracterizado como quase-experimental e a análise comparativa dos escores do instrumento são analisados na seção 7.2.

A variável independente foi a metodologia do PBL amparada pela TCC, aplicada a uma turma de Bioestatística para área da saúde para, envolvendo o conteúdo de THMP. Enquanto que, por variáveis dependentes, entendeu-se o grau de domínio do campo conceitual deste conteúdo, alcançado pelo aluno, após a aplicação do PBL, tomado a partir da análise de dados qualitativa (análise descritiva de duas questões abertas) e quantitativa (escore em um instrumento com dez itens fechados), caracterizando a abordagem metodológica mista entre qualitativa e quantitativa, por meio da triangulação metodológica.

No delineamento quase-experimental é preciso que o Pesquisador esteja ciente de quais variáveis específicas seu design particular falha em controlar Campbell e Stanley (2015). Neste sentido, não se pode afirmar que a variável independente seja, de fato, a única a interferir no resultado das variáveis dependentes, visto que questões subjetivas inerentes ao fazer pedagógico, independente da metodologia usada, podem ocorrer, como a motivação pela simples aprovação (busca da nota), a ansiedade estatística apontada na revisão de literatura (capítulo 2), a resistência à novas metodologias observada durante o estudo piloto (seção 5.1) entre outros fatores, alguns possivelmente detectáveis no diário de classe.

Na tentativa de estabelecer o nível de domínio do CC dos THMP melhora ao se utilizar um instrumento que leve em conta a expressão livre e escrita do aluno (abordagem quali e quantitativas), outro que seja mais focado em questões específicas do conteúdo visando o rendimento do aluno (abordagem quantitativa) e ainda um terceiro instrumento que permita a coleta da impressão do professor/tutor que orientou o processo (abordagem qualitativa).

Com isto, coaduna-se com a visão de [Augusto \(2014\)](#) e com o foco na triangulação.

Argumenta-se que, para um investigador, escolher uma metodologia de pesquisa não pode ser uma questão de preferência. Essa escolha terá de estar relacionada com as questões que o investigador coloca, com a natureza do que se pretende conhecer, com o tipo de respostas que espera providenciar. ([Augusto, 2014](#), p. 2)

A triangulação metodológica entre as abordagens qualitativa e quantitativa foi tal como define [Denzin \(1978\)](#),

Uma forma muito mais satisfatória de triangulação de métodos combina métodos dissimilares para medir a mesma unidade, o que chamo de between-methods ou across-methods. A justificativa para essa estratégia é que as falhas de um método geralmente são as forças de outro; e combinando métodos, os observadores podem alcançar o melhor de cada um, superando suas deficiências singulares. ([Denzin, 1978](#), p. 302)

Neste sentido, o instrumento para análise do nível de domínio (subseção [4.2.3](#)) contempla as duas abordagens de formas distintas e complementares. Primeiramente, as questões foram analisadas levando-se em conta a presença de invariantes operatórios esperados para as questões, separados em sete categorias e avaliados de acordo com uma pontuação (de zero a dois). A soma das pontuações gerou um escore final que serviu para caracterizar o nível de domínio do CC dos THMP, de forma quantitativa, entre baixo, moderado e alto. A segunda vertente do instrumento buscou analisar qualitativamente as respostas consideradas atípicas ou fora das categorias estabelecidas anteriormente.

As duas vertentes do instrumento promovem a complementariedade das abordagens qualitativa e quantitativa. A figura 4.2 ilustra esta abordagem.

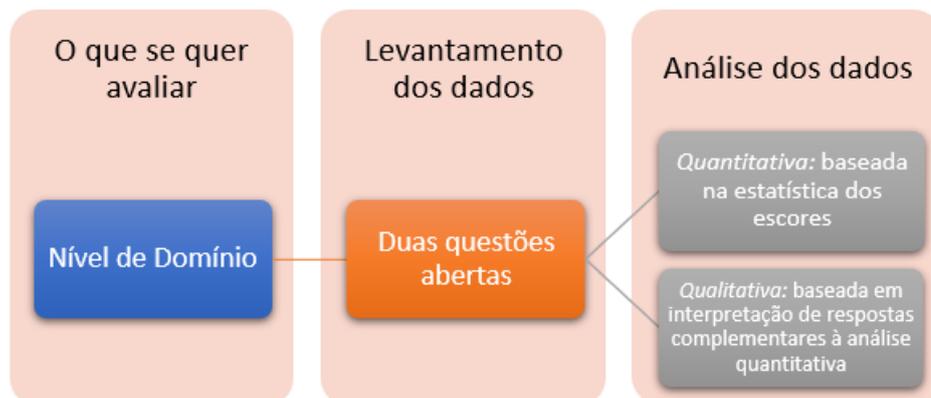


Figura 4.2: Desenho explicativo do instrumento para avaliar o nível de domínio do CC dos THMP.

Ainda destaca-se que o instrumento para medir o nível de domínio contempla a expressão escrita do aluno, onde seus esquemas de pensamento podem ficar mais aparentes, com a vantagem de se identificar possíveis invariantes operatórios, por ventura semelhantes, comuns entre os alunos da amostra.

De maneira conciliadora entre as abordagens qualitativas e quantitativas, [Gatti \(2004\)](#) pondera:

Os métodos de análise de dados que se traduzem por números podem ser muito úteis na compreensão de diversos problemas educacionais. Mais ainda, a combinação deste tipo de dados com dados oriundos de metodologias qualitativas, podem vir a enriquecer a compreensão de eventos, fatos, processos. As duas abordagens demandam, no entanto, o esforço de reflexão do pesquisador para dar sentido ao material levantado e analisado. ([Gatti, 2004](#), p. 13)

Quiz-se confrontar, de maneira complementar, as abordagens quali e quanti também no diário de classe e no instrumento fechado. O primeiro leva em conta os fatores que chamaram atenção do professor/pesquisador durante a aplicação do PBL, tanto no que

diz respeito ao desenvolvimento da metodologia quanto às características individuais externadas pelos Estudantes durante o processo. Por outro lado, o instrumento fechado avalia o rendimento estrito do estudante, tal como uma prova tradicional, típica da realidade dos estudantes da instituição de ensino superior (IES) onde a pesquisa ocorreu.

Faz sentido tomar-se como complementares o instrumento fechado e o diário de classe pois o primeiro representa a última instância avaliativa (fechada, objetiva e determinística) do processo descrito no diário de classe e que, em tese, foi determinante para o rendimento observado no instrumento fechado. Por esta abordagem, o diário de classe é a interpretação subjetiva de um processo que culminou na avaliação objetiva.

O instrumento fechado visa a avaliação do rendimento do estudante, com a vantagem de que as questões fechadas focam em conceitos específicos do CC dos THMP, que nem sempre aparecem naturalmente nas questões abertas. A figura 4.3 resume a ideia deste instrumento.



Figura 4.3: Desenho explicativo do instrumento para avaliar o rendimento dos alunos acerca dos THMP.

Isto posto, volta-se atenção para a amostra.

4.1 Contexto da Pesquisa e Caracterização da Amostra

A Instituição de Ensino Superior (IES) que recebeu as atividades desta Pesquisa está localizada no sul do Brasil, na cidade de Santa Maria, estado do Rio Grande do Sul.

A IES tem mais de 50 anos de experiência no ensino superior e conta com 35 cursos

de graduação, sendo oito da área da saúde (Biomedicina, Enfermagem, Fisioterapia, Medicina, Psicologia, Nutrição, Odontologia e Terapia Ocupacional), 27 especializações (dez na área da saúde), cinco mestrados (dois da área da saúde) e dois doutorados.

As atividades da Pesquisa ocorreram nos Cursos de Biomedicina, Psicologia e Nutrição, na disciplina de Bioestatística, lecionada por professores de Matemática, pois nenhum professor da IES é graduado em Estatística.

Quanto aos alunos ingressantes nestes cursos, seus rendimentos em matemática, na prova do vestibular dos anos de 2018 e 2019, são apresentados na tabela 4.1. A tabela mostra que o curso com melhor desempenho (Psicologia em 2019) alcançou 41,8% da nota total (5) e os demais rendimentos ficaram abaixo de 22,2% deste total, fazendo crer que os alunos ingressantes não trazem consigo uma forte bagagem matemática.

Tabela 4.1: Pontuação média em Matemática, de cinco pontos possíveis, na prova do vestibular de verão, dos estudantes ingressantes em 2018 e 2019, nos três cursos. Em 2019, a entrada para os alunos dos cursos de Biomedicina e Nutrição ocorreu por meio de redação. Fonte: Coordenadoria de Seleção e Ingresso da IES.

	Biomedicina	Nutrição	Psicologia
2018	1,11	0,96	1,08
2019	-	-	2,09

Os cursos em questão exibem uma descrição própria no sitio *web* da IES (transcritas no anexo B) que podem ser vistas como um perfil teórico dos egressos. Neste pretense perfil, apenas o curso de Biomedicina indica as ciências exatas como um de seus temas de estudo, o que pode indicar uma “afinidade” da Biomedicina com a Estatística, não explicitada pelos cursos de Nutrição e Psicologia, em suas descrições.

Fez-se uma busca na grade curricular dos cursos, disponíveis no sitio *web* da IES, a procura de disciplinas que tenham forte caráter quantitativo em suas ementas. O resultado passou pelo aval dos coordenadores dos cursos e pode ser visto a seguir.

Biomedicina Biofísica (1º semestre, 40 horas), Bioestatística (2º semestre, 40 horas)

e Epidemiologia (2º semestre, 40 horas).

Nutrição Técnica Dietética I¹ (2º semestre, 80 horas), Bioestatística (3º semestre, 60 horas) e Epidemiologia (3º semestre, 40 horas).

Psicologia Bioestatística (2º semestre, 60 horas).

Com isto, acredita-se que o curso de Biomedicina e Nutrição possuem uma inclinação maior para o aspecto quantitativo enquanto o curso de Psicologia só o leva em conta na disciplina de Bioestatística, foco desta pesquisa.

Procedeu-se, também, a uma pesquisa acerca das matrículas dos alunos amostrados, a fim de verificar se estes cursaram a disciplina de Bioestatística no tempo indicado. Esta informação pode vir do primeiro número da matrícula do aluno, que indica o ano de ingresso, juntamente com o segundo número, que refere-se ao semestre de entrada, naquele ano.

De posse destes dados, contabilizou-se quantos semestres o aluno está distante do seu semestre previsto para cursar a disciplina de Bioestatística, assim, o número zero é atribuído ao aluno que está fazendo a disciplina no período indicado, o número um indica um atraso de um semestre, o dois, dois semestres e assim por diante. O número menos um indica que o aluno fez a disciplina um semestre antes do tempo indicado.

A figura 4.4 traz as medidas resumo destas distâncias.

Pelos dados da 4.4, percebe-se que os alunos do curso de Psicologia costumam deixar de fazer a disciplina de Bioestatística no tempo determinado pelo curso. Isto pode indicar que os alunos estejam evitando a disciplina, o que corroboraria com a ideia de que o curso é avesso às questões quantitativas.

Indo além, foi feita uma busca nos trabalhos finais de graduação (TFG), defendidos no ano de 2019, em cada um dos três cursos, na intenção de identificar o uso de

¹Esta disciplina foi incluída pela Coordenadora do curso de Nutrição pois é nela que os cálculos nutricionais (que envolvem basicamente as proporções) são ensinados.

Turma		Estatística	
dist	NUT402	Média	,40
		Mediana	,00
		Desvio Padrão	1,265
		Mínimo	0
		Máximo	4
	PSC1010	Média	2,58
		Mediana	1,00
		Desvio Padrão	2,774
		Mínimo	0
		Máximo	9
	PSC445	Média	3,84
		Mediana	2,00
		Desvio Padrão	3,236
		Mínimo	-1
		Máximo	12

Figura 4.4: Medidas resumo, por turma, das distâncias do semestre no qual o aluno fez a disciplina de Bioestatística em relação ao semestre indicado pelo seu curso, em semestres, dos alunos que participaram do PBL. Gerado no SPSS.

ferramentas quantitativas nestes trabalhos, visto que a presença de tais ferramentas podem indicar uma inclinação do curso para a estatística.

A análise dos TFGs deu-se pela inspeção visual de todo o trabalho na busca de tabelas, gráficos ou números, seguida da leitura do título, do resumo e da metodologia de pesquisa (quando havia). Esta busca exploratória permitiu separar os trabalhos, quanto ao uso de abordagem quantitativa, como 'nenhum', 'pouco' e 'muito'.

A classificação 'nenhum', foi atribuída a estudos de cunho qualitativo e sem referência a métodos quantitativos. Já os classificados como 'pouco', trouxeram algum uso de métodos quantitativos como frequência absoluta ou porcentagens com ou sem suas representações gráficas (gráficos de colunas ou de setores). Por sua vez, os trabalhos que usaram significativamente métodos quantitativos foram classificados como 'muito' e apresentaram ampla análise descritiva, com uso para além de tabelas de frequência e porcentagens, como gráficos (histogramas, boxplots, etc), medidas de tendência central

ou estatística inferencial, por exemplo.

Os TFGs foram disponibilizados pela biblioteca da instituição, que não possuía todos os trabalhos do ano de 2019 à disposição, proporcionando um total de 53 trabalhos, dos quais 19 foram do curso de Biomedicina, apenas seis do curso de Nutrição e 28 do curso de Psicologia.

Após a análise, obteve-se a distribuição da tabela 4.2.

Tabela 4.2: Distribuição dos TFGs quanto ao grau de uso de análise quantitativa, nos três cursos.

	Biomedicina	Nutrição	Psicologia	Total
Nenhum	7	1	27	35
Pouco	4	2	1	7
Significativo	8	3	0	11
Total	19	6	28	53

Ao se juntar as categorias 'pouco' e 'muito', observa-se mais claramente o distanciamento do curso de Psicologia quanto ao uso de métodos quantitativos nos trabalhos finais de graduação, como se vê na tabela 4.3.

Tabela 4.3: Distribuição dos TFGs quanto ao uso ou não de análise quantitativa, nos três cursos.

	Biomedicina	Nutrição	Psicologia
Nenhum	37%	17%	96%
Pouco ou Muito	63%	83%	37%
Total	100%	100%	100%

Dos 99 estudantes da amostra, 61 participaram da metodologia designada por esta Pesquisa, destes, há ampla predominância de alunos do curso de Psicologia (ver figura 4.5), também por isto, as descrições feitas no capítulo 6 e 7 têm-se ao curso de

Psicologia.

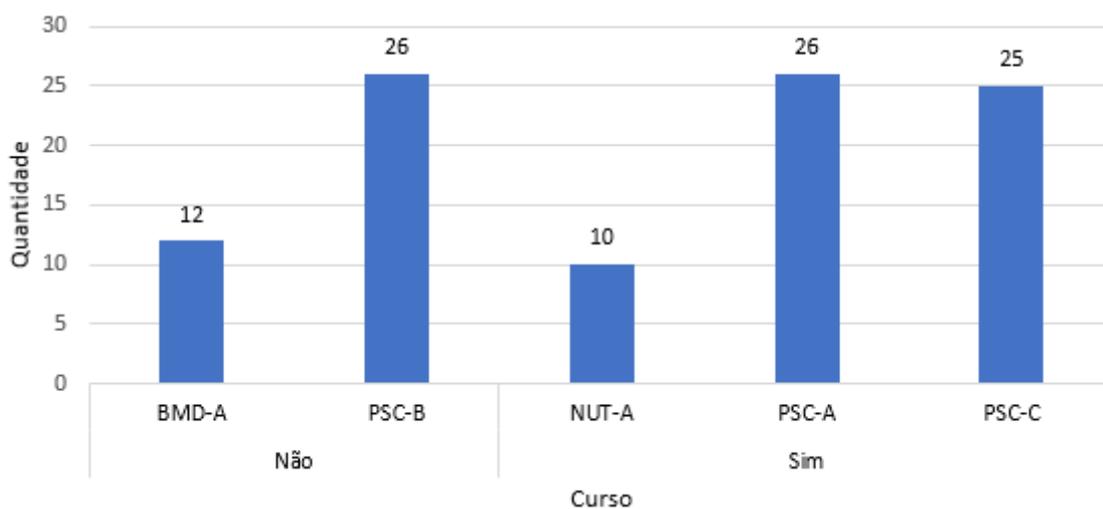


Figura 4.5: Distribuição dos alunos nos cursos que participaram do PBL (n = 61).

Considerando o baixo uso de processos quantitativos, leia-se, uma única disciplina com cunho quantitativo (Bioestatística), atraso dos alunos em fazer a disciplina no tempo indicado pelo curso e ínfima quantidade de trabalhos de final de graduação com cunho quantitativo, é de se esperar que os alunos venham a ter dificuldade com o uso de softwares como o Excel, por exemplo, que exigem um pensamento muito próximo ao matemático e que será utilizado nas aulas de Bioestatística.

É também passível de se esperar que haja grande ansiedade estatística nas turmas do curso de Psicologia, dada a grande incidência mostrada na revisão de literatura e pela característica de pouco uso de métodos quantitativos no curso.

Relativamente à disciplina onde se insere esta Pesquisa, a Bioestatística é a aplicação dos métodos estatísticos aos problemas biológicos e às ciências da saúde, como salientam [Callegari-Jacques \(2003\)](#) e [Pagano e Gauvreau \(2006\)](#). Para [Callegari-Jacques \(2003\)](#), a Bioestatística é um ramo da Estatística, visto que algumas técnicas estatísticas são mais empregadas no âmbito das ciências biológicas.

Nesta IES, a Bioestatística está dividida em três grandes áreas, como mostra a

figura 4.6:

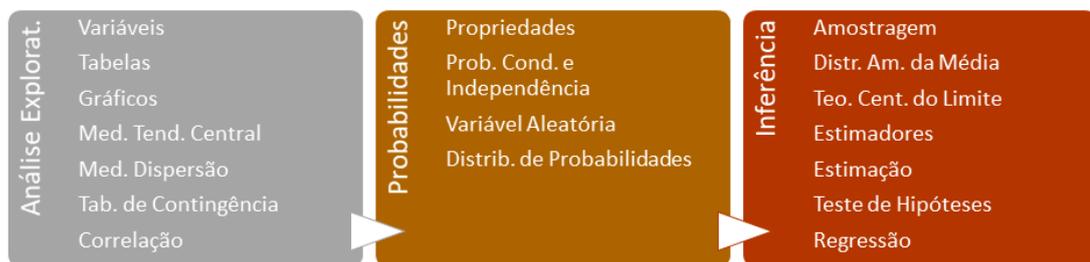


Figura 4.6: Estrutura da Estatística de um curso de graduação da área da saúde.

Os Testes de Hipóteses, conteúdo foco desta Pesquisa, estão dentro da área da Inferência Estatística, junto com a Estimação Estatística, e podem ser subdivididos como mostra a figura 4.7.

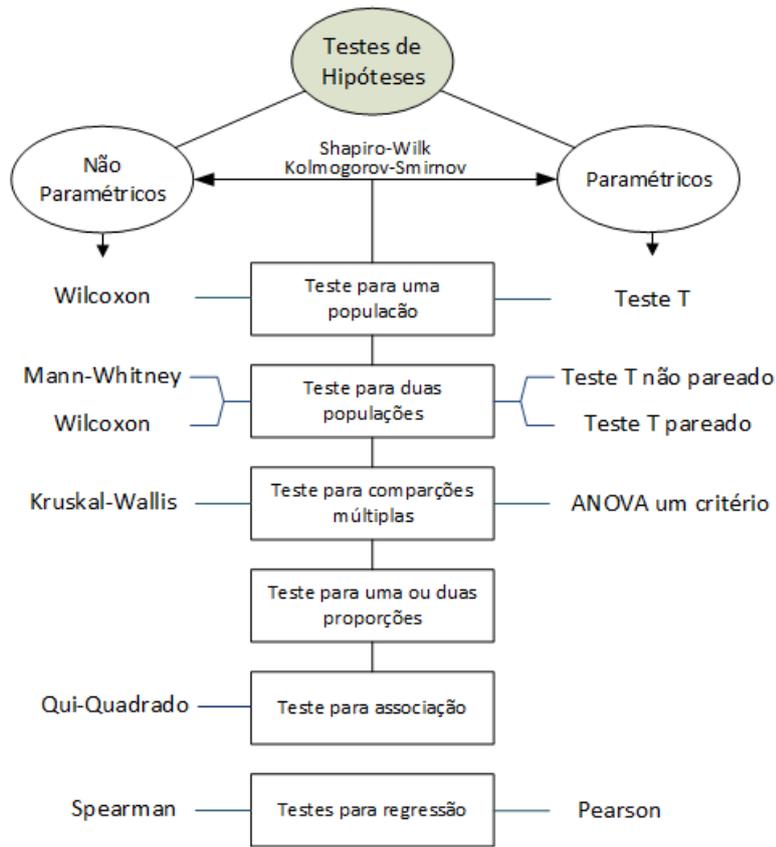


Figura 4.7: Testes de Hipóteses comumente trabalhados em um curso de Bioestatística para a graduação

Nesta IES, é obrigatória a construção, por parte do professor, de um documento norteador da disciplina no semestre onde está sendo ofertada, chamado Plano de Ensino. A disciplina de Bioestatística ministrada no escopo desta Pesquisa, considerou a seguinte ementa em seu Plano de Ensino:

- Papel da estatística na área da saúde.
- Organização da pesquisa.
- Medidas de posição e dispersão.
- Caracterização estatística de variáveis.

- Amostragem.
- Estimação estatística.
- Testes de hipóteses.
- Correlação e regressão.
- Testes não-paramétricos.
- Determinação do teste estatístico a utilizar.

Os conteúdos trabalhados foram propostos na seguinte ordem:

1. Estatística Descritiva
 - (a) Variáveis Estatísticas
 - (b) Tabelas e Gráficos
 - (c) Medidas Resumo
2. Fundamentos de Probabilidade
 - (a) Definição de Probabilidade
 - (b) Probabilidade Condicional
 - (c) Distribuição Normal
 - (d) Distribuição t
3. Teoria da Amostragem
 - (a) Tipos de Amostragem
 - (b) Distribuição Amostral da Média
 - (c) Teorema Central do Limite

4. Estimação de Parâmetros

- (a) Estimação da média
- (b) Estimação da proporção

5. Testes de Hipóteses

- (a) Hipótese Nula e Alternativa
- (b) Tipos de Erros
- (c) Nível de Significância e Poder do Teste
- (d) Teste para uma média populacional
- (e) p-valor

Dos conteúdos listados acima, os alunos que participaram do PBL, tiveram aulas expositivas acerca de alguns deles anteriores aos testes de hipóteses, quais sejam: Estatística Descritiva, Fundamentos de Probabilidade, Teoria da Amostragem e Estimação de Parâmetros.

A metodologia de ensino aqui chamada de PBL apoiada pela TCC (tratamento), foi aplicada a uma parte dos alunos enquanto os demais tiveram aulas tradicionais. Dentre 5 turmas (BMD-A, NUT-A, PSC-A, PSC-B e PSC-C), 3 foram escolhidas aleatoriamente para passarem pela metodologia proposta (NUT-A, PSC-A e PSC-C).

Os encontros presenciais aconteceram uma vez por semana, durante três períodos, totalizando 150 min semanais. Todos em laboratório de informática com acesso ao *softwares* Excel e Bioestat, além de navegadores para acesso à *internet*.

Desta forma, a disciplina de Bioestatística para os cursos de Psicologia e Nutrição contam com 60 horas semestrais, pouco ao se comparar com as 90 horas semestrais da disciplina de Estatística Descritiva e Geral², ministrada na Universidade Federal de

²Disponível em <http://coral.ufsm.br/estat/index.php/graduacao/2014-11-12-20-02-04/2014-11-12-20-10-05>

Santa Maria (UFSM) e que tem, praticamente, os mesmos conteúdos programáticos.

Mais detalhes acerca da caracterização da amostra e do contexto da Pesquisa estão na seção Materiais e Métodos, a seguir.

4.2 Materiais e Métodos

A seguir, passa-se a descrever os materiais (instrumentos) utilizados na Pesquisa, bem como os métodos utilizados para seu emprego.

4.2.1 Diário de Classe

O levantamento de dados qualitativos durante a aplicação do PBL foi feito por meio do diário de classe, que segundo [Fiorentini e Lorenzato \(2009\)](#) é um dos instrumentos mais ricos de coleta de informações. Os autores salientam que o diário de classe tem como objetivo registrar de maneira detalhada e sistematizada, os acontecimentos, as rotinas e as conversas que contribuirão no processo de análise das ocorrências observadas. Neste sentido, o diário pode conter uma dupla perspectiva, a descritiva e a interpretativa, que segundo os mesmos autores, convém que o pesquisador busque equilibra-las a fim de que o diário não seja meramente técnico ou muito genérico e superficial.

Na perspectiva de [Patton \(1980\)](#), os dados qualitativos se caracterizam por descrições detalhadas de fenômenos, comportamentos; citações diretas de pessoas sobre suas experiências; dados com maior riqueza de detalhes e profundidade e interações entre indivíduos e grupos, entre outras características.

[Fiorentini e Lorenzato \(2009\)](#) recomendam ao professor que pretende investigar sua própria prática, como no caso desta pesquisa, que ele elabore o diário logo após o término das aulas e lance mão de outros instrumentos de coleta de informações como gravações de áudio ou vídeo, coleta de registros dos alunos, entre outros.

O diário de classe auxiliará na detecção dos invariantes operatórios utilizados pelos

alunos e na adequação da atuação do professor na zona de desenvolvimento proximal destes estudantes a fim de bem evoluir no processo de domínio do campo conceitual durante o desenvolvimento da PBL.

Vergnaud ressalta a importância da observação do gesto da criança que, com o dedo apontando o seu objeto de contagem, diz em voz alta o número natural equivalente àquele objeto e, nesta análise, pode-se identificar invariantes operatórios Vergnaud (2017). Por este caminho, o diário de classe é auxiliar na identificação de sutilezas tando de um estudante em particular quanto de uma turma, em situação.

Em todas as aplicações o diário de classe teve seu registro feito em áudio para tornar a narrativa mais natural visto que é para o próprio professor, sem perder informação pela passagem do tempo, no caso da versão escrita. O diário não seguiu um roteiro de descrição para não restringir a análise da aula a pormenores desnecessários.

Entende-se que um diário de classe não é capaz de registrar todos os pormenores da sala de aula, entretanto, é uma ferramenta importante para o registro das impressões do professor/pesquisador, difíceis de observar com outros instrumentos.

4.2.2 Instrumento para Análise do Rendimento

No paradigma quantitativo, Fiorentini e Lorenzato (2009) sustentam que é de uso comum os experimentos didáticos com classes de controle e experimental, bem como a aplicação de testes e questionários fechados e padronizados onde os resultados são tratados estatisticamente tendo como parâmetro de análise uma escala validada previamente.

Com o objetivo de analisar o rendimento (score) de alunos da área da saúde que responderam a um instrumento fechado contendo itens típicos sobre os THMP, estratificados por PBL (participou ou não participou), por gênero (autodeclarados feminino e masculino) e turma (entre os cursos de Biomedicina (BMD), Nutrição (NUT) ou Psicologia (PSC)), desenvolveu-se uma AED seguida da análise de medidas típicas

da TCT e testes de hipóteses, descritas acima.

A metodologia de ensino aqui chamada de PBL apoiada pela TCC (tratamento), foi aplicada a uma parte dos alunos enquanto os demais tiveram aulas tradicionais. Dentre 5 turmas (BMD-A, NUT-A, PSC-A, PSC-B e PSC-C), 3 foram escolhidas aleatoriamente para passar pela metodologia proposta (NUT-A, PSC-A e PSC-C).

O instrumento foi aplicado a todos os estudantes. Todas as turmas são da área da saúde e consideradas semelhantes didaticamente. Neste sentido, o estudo pode ser caracterizado como quase-experimental e a análise comparativa dos escores do instrumento são analisados neste artigo.

O instrumento utilizado possui dez questões típicas, de múltipla escolha, tal como utilizadas em livros e materiais didáticos da área. Entretanto, mesmo questões típicas podem variar bastante entre professores e instituições, logo, o instrumento levou em conta tanto questões teóricas quanto práticas e as realidades de duas instituições de ensino superior (IES). Ele foi construído com o auxílio da Sigma Jr, empresa júnior do curso de Estatística da Universidade Federal de Santa Maria (UFSM) a qual gerou sua primeira versão consultando professores de estatística e bioestatística daquela universidade.

Em seguida, o questionário passou por avaliação de outros três professores da Universidade Franciscana (UFN), gerando algumas alterações antes de sua versão final, com 10 itens de múltipla escolha, que podem ser visualizados no anexo I.

O primeiro item, traz uma situação onde um profissional quer verificar se o conteúdo dos frascos de sabão líquido comprados por um hospital realmente tem a média indicada pelo fabricante, considerando sete amostras destes frascos. Os itens dois a seis, referem-se a uma mesma situação na qual se quer verificar se a pressão arterial média de crianças de uma certa comunidade é, de fato, maior do que a média nacional. Os itens restantes exploram questões conceituais acerca do nível de significância, p-valor e da estatística do teste.

		TURMA					
		BMD-A	NUT-A	PSC-A	PSC-B	PSC-C	Total
PBL	Não	12	0	0	26	0	38
	Sim	0	10	26	0	25	61
Total		12	10	26	26	25	99

Tabela 4.4: Quantidades de alunos por curso, turma e participação no PBL.

Os dados foram coletados nos dois semestres de 2018 e no primeiro semestre de 2019, nas turmas onde o professor pesquisador foi ministrante da disciplina de Bioestatística, distribuídos conforme a Tabela 4.4.

O instrumento foi aplicado a 99 estudantes de uma instituição de ensino superior da cidade de Santa Maria, RS, por meio do ambiente virtual de aprendizagem Moodle, de maneira presencial e individual, em laboratório de informática, durante duas horas/aula (100min) cada aplicação. Os estudantes tinham acesso à internet e eram familiarizados com os softwares Excel³ e Bioestat⁴.

Os dados coletados foram exportados do Moodle para uma planilha Excel onde foram devidamente limpos, restando as colunas Identificador, PBL, Curso, e a resposta certa (1) ou errada (0) em cada um dos 10 itens. Ainda na planilha Excel foi calculado o índice de discriminação.

A AED teve dois momentos distintos, um com foco no instrumento e no rendimento geral dos estudantes e o outro nos escores dentro das variáveis estratificadoras.

No primeiro momento lançou-se mão das medidas resumo dos escores proporcionadas pelo software IBM SPSS⁵, que também gerou o histograma com a curva normal subjacente, com visualização apoiada pelo boxplot feito pelo software Orange Canvas⁶. Em seguida, procedeu-se ao cálculo das medidas da TCT, por meio do pacote 'ltm' do software R, que gera também o gráfico dos escores vs proporção de acertos no item

³Planilha eletrônica da Microsoft, v. 2016.

⁴Software estatístico do Instituto Mamirauá, v. 5.4. Disponível em <https://www.mamiraua.org.br/downloads/programas/>

⁵v. 25

⁶v. 3.21.0

Rizopoulos (2018).

No segundo momento, cada comparação dentro das variáveis estratificadoras contou com boxplots comparativos, seguindo para a estimação de núcleo gaussiano, criados no Orange Canvas. Adiante, tem-se a comparação das médias dos grupos por meio de um gráfico de colunas gerado no software Tableau⁷ e testes de hipóteses.

Técnicas Estatísticas para Análise do Instrumento Fechado

A análise dos escores obtidos no instrumento de múltipla escolha (fechado) deu-se por meio da Análise Exploratória dos Dados (AED) e a Teoria Clássica dos Testes (TCT), entendidas como mostrado a seguir, além de testes de hipóteses para comparar os diferentes grupos.

Tukey (1977) pondera que a AED se refere ao olhar para os dados e ver o que eles parecem dizer, concentrando-se em aritmética simples e figuras fáceis de produzir. É preciso levar em conta que as aparências são descrições parciais que servem para se extrair novos *insights*, não se preocupando com a confirmação.

Batanero et al. (1991) lembram que antes de Tukey (1977) introduzir a AED, a análise se baseava fundamentalmente nos cálculos, trazendo duas consequências: a diminuição da importância visual da representação dos dados e a equiparação da análise a um modelo comprobatório preestabelecido, reduzindo a análise a um teste de hipóteses que barra a extração de qualquer outra informação que se possa deduzir dos dados. Neste sentido, Tukey (1977) apresenta o seguinte princípio: “É importante entender o que você pode fazer antes de aprender a medir o quão bem você parece ter feito”.

Para Behrens (1997), cada pesquisa desenvolve sua própria AED, segundo suas necessidades. Considerando este trabalho, a AED contou com visualização de dados, tabelas de frequências e medidas resumo.

Para além da AED, existe a TCT, que segundo Soares et al. (2018), é uma das vertentes da psicometria moderna, ela se preocupa em explicar o resultado final total

⁷v. 2019.2.3

(score) do teste. Neste trabalho apresenta-se a análise das medidas típicas encontradas na TCT para um instrumento (teste) específico, que são o índice de dificuldade, o índice de discriminação, o coeficiente bisserial e o coeficiente Alpha de Cronbach.

O índice de dificuldade ($Diff_i$) é dado pela proporção de respondentes que acertaram o item e mede o grau de dificuldade do item. Quanto maior for o índice de dificuldade, mais fácil é o item.

$$Diff_i = \frac{C_i}{T_i}$$

Onde, $C_i \rightarrow$ quantidade de respostas certas para o item i ;

$T_i \rightarrow$ total de respostas ao item i .

O índice de discriminação do item ($Disc_i$) é calculado por meio da diferença entre a proporção de acertos do escores mais extremos (escores maiores menos escores menores) e mede a capacidade do item de diferenciar respondentes de maior e menor habilidade. Quanto maior for o resultado do índice, maior será sua discriminação. “O grupo superior é constituído usualmente por 27% dos indivíduos que responderam ao teste: aqueles que obtiveram os maiores escores totais. O grupo inferior também é usualmente constituído por 27% dos indivíduos: aqueles que obtiveram os menores escores totais” (da Silveira, 2013).

$$Disc_i = E_i - e_i$$

Onde,

$E_i \rightarrow$ proporção de acertos no item i , dos 27% dos estudantes com maiores escores;

$e_i \rightarrow$ proporção de acertos no item i , dos 27% dos estudantes com menores escores.

Já o coeficiente bisserial mede o grau de associação entre a proporção de acertos do item e o valor do score na prova, ajudando a responder a seguinte questão: acertar um determinado item tem relação com escores altos na prova? O coeficiente bisserial,

pela notação de [Borgatto e de Andrade \(2012\)](#), é dado por:

$$r_{bis} = \frac{M^+ - M^-}{S} \cdot \frac{p(1-p)}{h(p)}$$

Onde,

M^+ → média do escore para os que acertaram o item;

M^- → média do escore para os que erraram o item;

S → desvio padrão do escore de todos os respondentes;

p → proporção de acerto do item;

$h(p)$ → valor da densidade da distribuição normal com média 0 e variância 1 no ponto em que a área da curva à esquerda deste ponto é igual a p .

Para o caso de não se poder presumir que os escores assumam uma distribuição normal, usa-se o coeficiente de correlação ponto-bisserial, dado a seguir:

$$r_{pbis} = \frac{M^+ - M^-}{S} \cdot \sqrt{\frac{p}{1-p}}$$

A fidedignidade do instrumento será auferida pelo coeficiente Alfa de Cronbach ([Cronbach, 1951](#)), que é dado por

$$\alpha = \left(\frac{k}{k-1} \right) \cdot \left(1 - \frac{\sum_{i=1}^k S_i^2}{S_t^2} \right)$$

Onde,

k → quantidade de itens do questionário;

S_i^2 → variância de cada item;

S_t^2 → variância total do escore.

O Alfa de Cronbach (α) mede a correlação entre as respostas do instrumento, que tem a mesma escala de medição, por meio da análise do perfil das respostas dadas pelos estudantes. Ele mede a fiabilidade do instrumento e quanto mais próximo de 1, mais

fiável é o instrumento, entretanto, não há consenso sobre o valor de α a partir do qual diz-se que o instrumento é altamente fiável (Maroco e Garcia-Marques, 2006).

A estimação não paramétrica de densidade por núcleo (kernel) com função de núcleo gaussiana será utilizada no contexto da AED e trata-se da estimação do formato da distribuição populacional por meio de uma suavização do histograma amostral buscando uma aproximação da curva normal que permite uma análise visual da distribuição dos dados (Estatcamp, 2019; Orange Data Mining, 2019).

Ainda na AED, o gráfico cartesiano que relaciona o escore com a proporção de acertos por item, permite verificar a evolução da proporção de acertos em cada item a medida em que o rendimento dos estudantes aumenta. O esforço dedicado à AED cria o ambiente propício para os testes de hipóteses, no sentido de cumprir com seus pré-requisitos e corroborar com seus resultados.

4.2.3 Instrumento para Análise do Nível de Domínio

O CC dos THMP foi estabelecido em 3.4.1 e o nível de domínio que um estudante alcançou depois de ter trabalhado seus conceitos pode ser melhor entendido por meio da aplicação de um instrumento que leve em conta fatores relevantes deste CC em suas situações típicas.

Estes fatores relevantes são compostos por invariantes operatórios, cujo esforço da detecção pode ser auxiliado com o uso do instrumento para a aferição do nível de domínio do CC dos THMP, que leve em conta as respostas esperadas a duas questões abertas, aplicadas à amostra ao final do processo de PBL.

O instrumento procura detectar, por meio da análise de duas questões acerca dos THMP, invariantes operatórios constituintes de sete categorias necessárias para o desenvolvimento satisfatório de tais questões, que são: Hipótese Nula, Hipótese Alternativa, Nível de Significância, Estatística do Teste, P-valor, Decisão e Interpretação.

Desta forma, os conceitos fundamentais dos THMP, considerando as mudanças

propostas no estudo piloto (5.1), estão preservados. Quais sejam:

1. Determinação das hipóteses nula e alternativa;
2. Estabelecimento do nível de significância (α);
3. Cálculo da estatística do teste e/ou do p-valor;
4. Tomada de decisão e interpretação do resultado do teste.

O processo de levantamento de dados a partir das respostas às questões 1 e 2, por meio deste instrumento, consistiu na leitura da resolução da questão 1 e 2 de cada aluno, buscando a resposta à categoria “Hipótese Nula” e atribuindo os valores 0, 1 ou 2 à esta resposta conforme a tabela 4.5, onde a coluna “Categoria” refere-se à categoria de análise baseada no Campo Conceitual descrito na seção 3.4. Assim, analisou-se a categoria “Hipótese Nula” de todos os alunos, nas duas questões. Em seguida, passou-se para as demais categorias, na ordem apresentada na tabela 4.5, seguindo a mesma metodologia de levantamento de dados.

Alguns comentários sobre algumas descrições (terceira coluna) da tabela 4.5:

- O nível de significância precisa ser definido antes da comparação com p-valor. Entretanto, na questão 1, ele já estava informado no corpo da questão, estando previamente definido.
- O *software* Bioestat obriga o usuário a escolher a estatística do teste antes de realizar os cálculos. Isto aparece de forma implícita na resolução das questões e pode ser identificado pelo valor da estatística ou do p-valor que o aluno usou.
- A etapa da decisão refere-se à hipótese nula e alternativa, ou seja, a hipótese alternativa precisa ser evidenciada quando há a rejeição da hipótese nula.

Os dados levantados foram colocados em planilha Excel, tal como mostrado na figura ??, e ecores foram calculados, de tal forma que a soma das pontuações dos alunos

Tabela 4.5: Tabela contendo a categoria de análise, o escore atribuído e a descrição de cada nível.

Categoria	Nível (Escore)	Descrição
Hipótese Nula	Não apresenta/ Não operacional (0)	H_0 não é referida implícita ou explicitamente ou é referida sem ser operacional. ex.: $H_0 : \mu = 1,3$ onde o correto é $H_0 : \mu = 1,4$ ou $H_0 : \mu \neq 1,4$ e $H_1 : \mu \neq 1,4$
	Domina/Explicita Parcialmente (1)	H_0 é explícita não totalmente correta, por expressão matemática ou de forma cursiva. ex.: $H_0 = 1,4$ ou $H_0 : \mu = \mu_0$ sem dizer quanto vale μ_0
	Domina/Explicita Totalmente (2)	H_0 é explícita total e corretamente, por expressão matemática ou de forma cursiva. ex.: $H_0 : \mu = 1,4$
Hipótese Alternativa	Não apresenta/ Não operacional (0)	H_1 não é referida implícita ou explicitamente ou é referida sem ser operacional. ex.: $H_1 : \mu \neq 1,3$ onde o correto é $H_1 : \mu \neq 1,4$
	Domina/Explicita Parcialmente (1)	H_1 é explícita não totalmente correta, por expressão matemática ou de forma cursiva. ex.: $H_1 \neq 1,4$ ou $H_1 : \mu \neq \mu_0$ sem dizer quanto vale μ_0
	Domina/Explicita Totalmente (2)	H_1 é explícita total e corretamente, por expressão matemática ou de forma cursiva. ex.: $H_1 : \mu \neq 1,4$
Nível de Significância	Não apresenta/ Não operacional (0)	α não é referido implícita ou explicitamente ou é referido sem ser operacional. ex.: $\alpha = 0,01$ onde o correto é $\alpha = 0,05$
	Domina/Explicita Parcialmente (1)	α é explícito não totalmente correto, por expressão matemática ou de forma cursiva. ex.: α é utilizado somente na comparação com p-valor, sem ser estabelecido antes do cálculo do p-valor.
	Domina/Explicita Totalmente (2)	α é explícito total e corretamente, por expressão matemática ou de forma cursiva. ex.: $\alpha = 0,05$
Estatística do Teste (T ou Z)	Não apresenta/ Não operacional (0)	Z ou T não é referido implícita ou explicitamente ou é referido sem ser operacional. ex.: T onde o correto é Z
	Domina/Explicita Parcialmente (1)	Z ou T é explícito ou implícito não totalmente correto, por expressão matemática, gráfico ou de forma cursiva. ex.: a região marcada no gráfico é incorreta, ou cálculo de z ou t incorretos ou escolha de Z ou T não justificada.
	Domina/Explicita Totalmente (2)	Z ou T é explícito ou implícito total e corretamente, por expressão matemática, gráfico ou de forma cursiva. ex.: Z ou T escolhidos e calculados apropriadamente
P-valor ou cálculo de t ou z	Não apresenta/ Não operacional (0)	p-valor ou o cálculo da estatística do teste não é referido implícita ou explicitamente ou é referido sem ser operacional.ex.: escolheu-se o p bilateral quando o correto era o unilateral ou houve a escolha errada da estatística do teste.
	Domina/Explicita Parcialmente (1)	p ou cálculo da estatística do teste é explícito não totalmente correto, por expressão matemática, gráfico ou de forma cursiva. ex.: erro nas casas decimais do p-valor que não altera a decisão ou escolha correta mas cálculo incorreto de t ou z
	Domina/Explicita Totalmente (2)	p-valor ou cálculo da estatística do teste é explícito total e corretamente, por expressão matemática, gráfico ou de forma cursiva. ex.: $p = 0,0017$

Tabela 4.6: Tabela contendo a categoria de análise, o escore atribuído e a descrição de cada nível (cont.).

Categoria	Nível (Escore)	Descrição
Decisão	Não apresenta/ Não operacional (0)	Não há tomada de decisão no teste ou a decisão não é operacional.
	Domina/Explícita Parcialmente (1)	Decisão é explícita não totalmente correta ou incompleta, por expressão matemática, gráfico ou de forma cursiva. ex.: $p < \alpha$ logo aceita-se H_0 ou $p < \alpha$ logo, rejeita-se H_0
	Domina/Explícita Totalmente (2)	Decisão é explícita total e corretamente, por expressão matemática, gráfico ou de forma cursiva. ex.: $p < \alpha$ logo, rejeita-se H_0 em benefício de H_1 .
Interpretação	Não apresenta/ Não operacional (0)	Interpretação não é referido implícita ou explicitamente ou é referida sem ser operacional. ex.: a amostra provém da população em questão considerando sua média e desvio padrão.
	Domina/Explícita Parcialmente (1)	Interpretação é explícita não totalmente correta, por expressão matemática ou de forma cursiva. ex.: Não é provável retirar uma amostra com média 1,3 de uma população com média 1,4.
	Domina/Explícita Totalmente (2)	Interpretação é explícita total e corretamente, por expressão matemática ou de forma cursiva. ex.: O teste não sustenta a ideia de que a amostra tenha vindo da população em questão.

nas diferentes categorias gerou um escore da questão (SCORE Q1 e SCORE Q2) e a soma dos escores de cada questão gerou o escore total do aluno referente às duas questões (SCORE TOTAL = SCORE Q1 + SCORE Q2).

Toda a análise foi estratificada pela participação ou não no PBL (PBL: não (0) ou sim (1)), a fim de se conhecer as diferenças que a metodologia gerou na amostra.

Os escores passaram por uma análise descritiva contendo média e mediana, valor mínimo e máximo e amplitude, amplitude interquartil e desvio padrão.

Para a análise gráfica utilizou-se *boxplots* e histogramas. Os *boxplots* apresentam média, quartis e *whiskers* e os histogramas foram gerados respeitando as classes da tabela 7.4.

A comparação entre os grupos que participaram ou não do PBL deu-se pela comparação dos respectivos escores por meio do teste de T ou Mann-Whitney, dependendo da suposição de normalidade populacional indicada pelo teste de Shapiro-Wilk. Todos os

Tabela 4.7: Tabela de classificação do nível de domínio do CC dos THMP de acordo com o escore total no instrumento para questões abertas.

Escore Total	Nível de Domínio
(0,10]	Baixo
(10,20]	Moderado
(20,30]	Ampla

testes levaram em conta um nível de significância de 5%. Os resultados desta análise estão na subseção 7.3.1.

Uma classificação do nível de apropriação do aluno em relação ao CC dos THMP pode ser estabelecido com a tabela 7.4.

A seguir, apresenta-se uma resposta esperada a cada uma das duas questões do instrumento, considerando o CC dos THMP definido na seção 3.4.1 bem como as ponderações feitas a partir dos resultados do estudo piloto, visíveis na seção 5.1.1.

As questões 1 e 2 foram retiradas e adaptadas de dois livros de Bioestatística, Callegari-Jacques (2003, p. 188) e Díaz e López (2012, p. 204), respectivamente, e foram escolhidas por envolverem as sete categorias de análise do instrumento aqui apresentado, como se quer mostrar na descrição a seguir.

1. *A taxa de fenilalanina no soro, em indivíduos normais, é de 1,4 mg/100mL e o desvio padrão é de 0,32 mg/100mL. Supõe-se que cem pacientes formam uma amostra aleatória da população de indivíduos normais, com média amostral de 1,3 mg/100mL.*

Considerando um nível de significância de 1%, os dados estão de acordo com a suposição feita?

O exercício considera uma amostra que supostamente provém de uma população com média 1,4 mg/100mL e desvio padrão 0,32 mg/100mL. O que se quer é verificar a plausibilidade de uma amostra com 100 indivíduos e com média de 1,3 mg/100mL sair da população em questão, a um nível de significância de 1%. Esta

abordagem é típica dos THMP, devendo-se aplicar seu método de resolução para dar conta do exercício. Saber identificar que a situação refere-se ao CC dos THMP é um importante, mas nem sempre fácil, passo da resolução. No caso destes dois exercícios esta dificuldade não aparece visto que foram aplicados como avaliação dos THMP.

O entendimento total do exercício, seus dados, objetivos, processo de cálculo a serem utilizado, compõem os conceitos em ação necessários para a resolução do exercício.

A resolução inicia com o estabelecimento das hipóteses nula e alternativa. Na hipótese nula, o aluno precisa reconhecer que o valor de referência é a média populacional:

$$H_0 : \mu = 1,4mg/mL$$

Conceito em ação requerido: estabelecimento correto da hipótese nula.

Subsequentemente à hipótese nula, a hipótese alternativa deve ser estabelecida conforme o que indica a situação de maneira a contrastar a hipótese nula. Neste caso, a hipótese alternativa contrapõe a nula com a afirmação da diferença:

$$H_1 : \mu \neq 1,4mg/mL$$

Pela hipótese alternativa estabelecida, tem-se um teste bicaudal.

Conceito em ação requerido: estabelecimento correto da hipótese alternativa e identificação do tipo de teste.

A determinação do nível de confiança costuma vir na sequência, mas neste caso, ele já foi estabelecido: $\alpha = 1\%$.

Conceito em ação requerido: estabelecimento ou identificação do nível de confiança a ser utilizado.

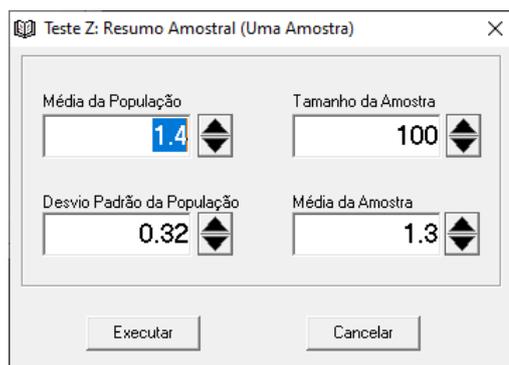


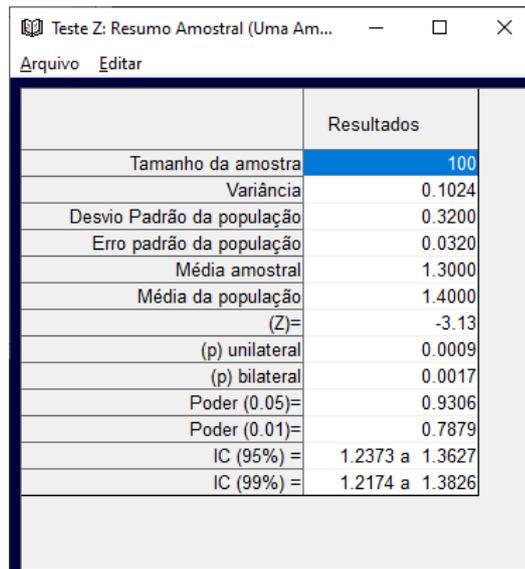
Figura 4.8: Entrada do Bioestat dos valores para Exercício 1. Acessível por: Estatísticas > Uma Amostra > Teste Z: Resumo Amostral.

A esta altura, surge a particularidade da abordagem escolhida para o THMP, a do p-valor ou a da estatística do teste. Como o estudo piloto indicou a necessidade de uma abordagem mais pragmática em vista da falta de tempo disponível, o que foi feito, e dado que os alunos responderam à estas questões em laboratório de informática munidos do software estatístico Bioestat, é esperado que utilizem a abordagem do p-valor. Não que elas sejam dissociadas conceitualmente, mas são distintas no que se refere ao processo de cálculo usando software.

Neste caso, o cálculo do p-valor é feito selecionando-se o local adequado dentro do software e introduzindo corretamente os valores pedidos. Deve se levar em conta que existe a pré-suposição de normalidade da distribuição populacional, que o desvio padrão populacional é conhecido e que a amostra é grande (maior do que 30 indivíduos), logo, o teste a ser utilizado é o Z, cujo preenchimento no Bioestat se vê na figura 4.8.

Após clicar em Executar, mostrado na figura 4.8, a saída correspondente pode ser visualizada na figura 4.9.

Neste ponto, espera-se que o aluno selecione o p-valor bilateral (0,0017) e o compare com o nível de significância $\alpha = 0,01$, determinando que o p-valor é



	Resultados
Tamanho da amostra	100
Variância	0.1024
Desvio Padrão da população	0.3200
Erro padrão da população	0.0320
Média amostral	1.3000
Média da população	1.4000
(Z)=	-3.13
(p) unilateral	0.0009
(p) bilateral	0.0017
Poder (0.05)=	0.9306
Poder (0.01)=	0.7879
IC (95%) =	1.2373 a 1.3627
IC (99%) =	1.2174 a 1.3826

Figura 4.9: Saída do Bioestat para o Exercício 1.

menor do que o nível de significância estabelecido, levando à decisão de rejeitar a hipótese nula em benefício da hipótese alternativa.

Teorema em ação requerido: se $p \leq \alpha$ então rejeita H_0 e aceita H_1 , ou, se $p > \alpha$ então não rejeita H_0 .

A tomada de decisão do THMP não é a sua interpretação, que pode ser expressa pela frase: não é plausível, a um nível de significância de 1% que a amostra provenha da população em questão. Porém, é desejável que a interpretação refira-se ao contexto da questão, como por exemplo: é baixa a probabilidade de se retirar uma amostra de 100 indivíduos com média amostral de 1,3 mg/100mL de uma população com média de 1,4 mg/100mL e desvio padrão de 0,32 mg/100mL de fenilalanina no soro.

2. A média de acidentes mortais em uma cidade é de 12 mensais. Após uma campanha de conscientização sobre os perigos no trânsito, nos seis meses consecutivos contabilizaram-se 8, 11, 9, 7, 10 e 9 acidentes mortais. A campanha foi efetiva?

Trata-se da comparação entre uma amostra com seis elementos e uma população com média igual a 12. Pela perspectiva dos TH, se quer verificar se a média amostral observada difere significativamente da média populacional indicada, logo trata-se de uma situação de THMP, que inicia com a definição da hipótese nula.

$$H_0 : \mu = 12$$

Conceito em ação requerido: estabelecimento correto da hipótese nula.

O exercício pergunta se a campanha foi efetiva o que pode suscitar dois tipos de hipótese alternativa: uma que indica que $\mu \neq 12$, sem uma pré-suposição sobre os valores amostrais ou outra afirmando que $\mu > 12$ pré-supondo que a campanha foi positiva.

Conceito em ação requerido: estabelecimento correto da hipótese alternativa e identificação do tipo de teste.

Comumente, a próxima ação é a determinação do nível de significância, que não tem indicação no exercício, deixando em aberto esta decisão.

Conceito em ação requerido: estabelecimento ou identificação do nível de confiança a ser utilizado.

A próxima etapa, considerando alunos com acesso a software estatístico, é a determinação da estatística do teste ou do p-valor. Pela praticidade, é esperado que os alunos determinem o p-valor.

Considerando que a amostra é pequena (seis elementos) e que se possui seus valores, utiliza-se o teste t. No *software* Bioestat, digita-se os dados no *grid*, seleciona-se o caminho para teste t: dados amostrais, insere-se a média populacional para obter-se a saída mostrada na figura 4.10.

Dependendo da hipótese alternativa escolhida, o p-valor unilateral (0,0017) ou

- 1 -	
Tamanho da Amostra	6
Média da População	12,0000
Média Amostral	9,0000
Erro Padrão	0,5774
(t)=	-5,1962
Graus de liberdade	5
(p) unilateral =	0,0017
(p) bilateral =	0,0035
Poder (0.05)	0,9998
Poder (0.01)	0,9979
IC 95% (média amostral)=	7,5156 a 10,4844
IC 99% (média amostral)=	6,6721 a 11,3279

Figura 4.10: Saída do Bioestat dos valores para Exercício 2. Acessível por: Estatísticas > Uma Amostra > Teste t: Dados Amostrais.

bilateral (0,0035) deve ser escolhido e comparado com o nível de significância estabelecido.

Teorema em ação requerido: se $p \leq \alpha$ então rejeita H_0 e aceita H_1 , ou, se $p > \alpha$ então não rejeita H_0 .

A interpretação do exercício pode ser a seguinte: a média de acidentes nos seis meses posteriores (9 acidentes por mês) foi significativamente menor do que a média histórica (12 acidentes por mês) ao nível de significância estabelecido, assim, o que corrobora com uma avaliação positiva da campanha conduzida.

Finalmente, cabe informar que este instrumento para a análise do nível de domínio do campo conceitual dos testes de hipóteses para a média populacional passou pela validação de dois Professores de Bioestatística da instituição de ensino superior onde esta Pesquisa se desenvolveu.

Um dos Professores é doutor em bioestatística e leciona esta disciplina para a área da saúde a quinze anos, além de fazer parte do programa de pós-graduação em ensino de ciências da mesma instituição.

O outro professor é doutor em ensino de ciências com pesquisas acerca da Teoria dos Campos Conceituais e leciona a disciplina de bioestatística a quatro anos.

Ambos deram o parecer favorável ao uso do instrumento, porém, o primeiro professor sugeriu que o instrumento contivesse mais questões e o segundo, que em construção futura, as situações sejam mais variadas.

Neste sentido, ressaltamos que o aumento do tamanho do instrumento inviabiliza a análise pormenorizada para uma amostra de 99 Estudantes.

Capítulo 5

Atividades Prévias

Foram feitas três aplicações da metodologia do PBL com a TCC, a primeira, a título de estudo piloto, descrito abaixo (5.1), que revelou a necessidade de ajustes na metodologia. O estudo piloto também mostrou que era preciso que os alunos experimentassem a metodologia mais cedo durante a disciplina, o que gerou a segunda aplicação (5.2) acerca do conteúdo de dispersão. Finalmente, houve a aplicação definitiva do PBL com TCC (6), que gerou os dados para esta Tese.

5.1 Estudo Piloto

Com o objetivo de aprimorar o processo investigativo, foi desenvolvido um estudo piloto, que segundo [Canhota \(2008\)](#), ajuda a elaborar um protocolo de investigação e a antever o resultado de um protocolo já elaborado, mas ainda não utilizado.

O estudo piloto ocorreu no primeiro semestre de 2018, com duas turmas do curso de Odontologia (20 alunos em cada turma), na disciplina de Bioestatística. O currículo da disciplina se estendia por 17 semanas, cada semana com dois encontros em laboratório de informática, cada encontro com duas horas/aula de 50 min cada aula. As atividades envolvendo o PBL foram aplicadas durante a 11^a e 16^a semanas.

Segue a descrição das atividades desenvolvidas no estudo piloto, e a análise de seus resultados, com base no diário de classe do professor/tutor.

No primeiro encontro, os grupos receberam, pelo Moodle, a atividade transcrita abaixo¹.

TESTE DE HIPÓTESES

Através do PBL (Problem-Based Learning)

O uso dos intervalos de confiança representa uma forma bastante informativa de apresentar os achados principais de um estudo que faz uso de amostras, como no caso da estimação da média, em que esses resultados são extrapolados para a população que originou os dados. Em alguns estudos, pode haver o objetivo de testar hipóteses formuladas sobre parâmetros de interesse. O teste de hipótese é um procedimento de inferência estatística usual nesses casos. Intervalos de confiança e testes de hipóteses estão intimamente relacionados, dado que seus fundamentos teóricos, em grande parte, são os mesmos. Entretanto, os testes de hipóteses têm algumas características importantes quanto à interpretação de seus resultados.

Uma hipótese estatística é uma suposição sobre um parâmetro de interesse e um teste estatístico de hipóteses é uma regra utilizada para decidir quando rejeitar uma hipótese, sendo essa regra sempre baseada em uma amostra aleatória.

Situação Problema

O arquivo “Diabete mulheres 21+” contém uma série de variáveis com dados extraídos de uma amostra aleatória de mulheres maiores de 21 anos.

Os pesquisadores se depararam com importantes questões acerca da

¹A transcrição foi feita com outra fonte e texto alinhado à esquerda com o intuito de remeter o leitor ao ambiente Moodle. As capturas de telas do Moodle podem ser vistas no anexo [D](#).

população das mulheres com mais de 21 anos:

Considerando que o nível de glicose dito normal deve estar abaixo de 99 mg/dl, a amostra que temos, permite aceitar a hipótese de que as mulheres desta população tem uma glicose normal?

Outras perguntas surgem desta, por exemplo:

Qual a hipótese pode ser formulada nesta situação? A amostra permite aceitar ou rejeitar esta hipótese? Sob quais condições se aceita ou rejeita a hipótese?

Esperava-se que os alunos identificassem e classificassem a variável a ser estudada (Glicose: coluna B da figura 5.1) em meio as demais, que calculassem a sua média (120,85 mg/dL), que a comparassem com o valor de referência(99 mg/dL) e fizessem ponderações a respeito, que permitiriam expandir a discussão para o uso das medidas de variabilidade, da relação entre população e amostra, do comportamento de amostras (distribuição amostral), do uso de intervalo de confiança, formulação de hipóteses, etc. Tais questões pautariam a pesquisa individual de cada integrante do grupo e seus achados seriam compartilhados no segundo encontro, dando seguimento ao processo do PBL.

	A	B	C	D	E	F	G	H	I
1	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
2	6	148	72	35	0	33,6	0,627	50	1
3	1	85	66	29	0	26,6	0,351	31	0
4	8	183	64	0	0	23,3	0,672	32	1
5	1	89	66	23	94	28,1	0,167	21	0
6	0	137	40	35	168	43,1	2,288	33	1
7	5	116	74	0	0	25,6	0,201	30	0
8	3	78	50	32	88	31	0,248	26	1
9	10	115	0	0	0	35,3	0,134	29	0
10	2	197	70	45	543	30,5	0,158	53	1
11	8	125	96	0	0	0	0,232	54	1
12	4	110	92	0	0	37,6	0,191	30	0

Figura 5.1: Cabeçalho e primeiros dados da base de dados do estudo piloto, referido com “Diabete mulheres 21+”.

Entretanto, os alunos não entenderam que as outras variáveis da base de dados não

estavam sob análise, levando o professor a parar a atividade e esclarecer que era preciso ater-se à variável glicose.

Voltando a trabalharem nos grupos, os alunos não produziram novas questões norteadoras, mas procuraram responder rapidamente somente às questões propostas de maneira simples e direta, como por exemplo, “a média da glicose é 120,85 mg/dL, logo as mulheres estão com a glicose acima do normal”. O questionamento “É só isso que tem pra fazer, professor?” foi recorrente.

Então, o professor fez o seguinte questionamento “A média da amostra é de 120,85 mg/dL, maior do que os 99 mg/dL indicados. Será que esta diferença é grande o suficiente para afirmar que a população de onde veio esta amostra tem a glicose elevada, ou é preciso considerar a variabilidade dos dados amostrais?” O questionamento não os incitou a continuarem pesquisando, nem mesmo a pensarem que estavam diante de um problema, apenas alguns perguntaram: “O Senhor quer que calcule o desvio padrão?”

Considerando que a atividade foi subsequente ao conteúdo de estimação de parâmetros e que este foi citado no texto da atividade, é notável que ele não tenha sido usado pelos grupos.

O processo continuou sob forte coordenação do professor que orientou grupo a grupo a tentarem abordagens conhecidas por eles como o desenvolvimento de uma estatística descritiva e gráficos, bem como ponderações sobre a relação entre população e amostra ou estimação de parâmetros e, a partir daí, tentarem ir para a formulação de hipóteses e procurarem o que são os testes de hipóteses.

Ao final do encontro, acordou-se que os grupos continuariam a pensar nas questões norteadoras e as trariam no próximo encontro.

Os THMP são técnicas desenvolvidas ao longo de muitos anos por muitos estatísticos experientes. O estudo piloto deixou claro que dificilmente os alunos se quer arranharão a técnica dos THMP simplesmente analisando a situação proposta, mesmo que esta seja típica da abordagem dos THMP.

Atendendo a esta demanda e considerando que o processo do primeiro encontro transcorreu muito mais lentamente do que o esperado e que os alunos foram muito menos diligentes do que o necessário, o professor/tutor disponibilizou, no Moodle, uma série de materiais sobre testes de hipóteses: um vídeo intitulado “O que são Testes de Hipóteses?²”, outro vídeo sobre para que servem os Testes de Hipóteses³ e um arquivo em pdf sobre as Hipóteses Nula e Alternativa, além do Manual do *software* Bioestat no Moodle, apresentados tal como se vê no apêndice D e seguido do seguinte texto.

Este material é importante para complementar a aula anterior e responder a algumas questões levantadas, além de preparar a próxima aula.

Material extraído do livro Bioestatística para os Cursos de Graduação da Área da Saúde, de Edson Zangiacomi Martinez.

No segundo encontro, os alunos não trouxeram questões norteadoras, mas dúvidas sobre o que era para fazer. Houve uma enxurrada de perguntas acerca de como os testes de hipóteses funcionavam e qual sua relevância naquela situação, mas sempre no sentido de tarefa, de rotina ou procedimento.

Também ficou evidente a insegurança dos alunos com o processo, principalmente com o fato de “não terem tido aula sobre testes de hipóteses”, referindo-se à aula expositiva. Também houve afirmações do tipo “Não entendi nada daqueles vídeos” ou “É impossível aprender aquilo tudo sozinho” ou ainda “Se o Senhor não vai dar aula, como vai cobrar na prova?”

Diante disto, o professor propôs às turmas algumas questões norteadoras condizentes com algumas de suas inquietações e com potencialidade de desenvolvimento do conteúdo, como por exemplo:

O que é uma hipótese estatística?

²Acessível, até maio de 2020, em <https://youtu.be/FZHqrVyc104>

³Acessível, até maio de 2020, em <https://youtu.be/h4QcWDD1rW0>

Considerando a hipótese verdadeira, é plausível obter a amostra com as características que se obteve?

Se fossem retiradas outras amostras, os resultados seriam todos semelhantes?

Ao se retirarem muitas amostras, como suas médias se comportariam?

Qual a probabilidade de se retirar uma amostra com as características da amostra em análise?

A discussão acerca destas questões transcorreram durante o restante do encontro e ficou acordado que os grupos trariam suas respostas a estas perguntas e suas perguntas nortedoras, por escrito, no terceiro encontro.

No terceiro encontro, houve o debate com a turma sobre as questões propostas acima. As questões se mostraram promissoras na geração de debates. Com elas foi possível explorar assuntos relevantes acerca dos fundamentos dos Testes de Hipóteses, porém, o processo de cálculo do THMP ainda estava longe de se estabelecer.

A fim de atender à demanda gerada no terceiro encontro, o professor fez um vídeo com a estrutura dos THMP e disponibilizou aos alunos⁴, via Moodle, tal como pode ser visto no anexo D).

No quarto encontro, os alunos que haviam estudado o vídeo, foram mais participativos durante a formalização do processo de cálculo dos THMP trabalhado de forma expositiva e dialogada pelo professor. Na segunda parte deste encontro, foi apresentado aos alunos o processo de cálculo usando o *software* Bioestat. Neste ponto, vale ressaltar que o processo de formalização dos conceitos cabe ao professor/tutor, dado que é durante este fechamento teórico que arestas são aparadas, processos são melhor definidos, entendidos, aplicados, formalizados, etc. Os Alunos tendem a ficar perdidos durante o processo do PBL no sentido de que não têm pleno conhecimento sobre o que estão fazendo e o fechamento dado pelo professor/tutor dá a segurança teórica que eles precisam, levando em conta todo o ganho que tiveram em suas pesquisas.

⁴Disponível em <https://youtu.be/hPdxUbNrMqE>

No quinto encontro, estando o processo de cálculo e as bases teóricas melhor estabelecidas, novas situações problemas foram introduzidas (Situações 2 a 5 mostradas em H). Os grupos trabalharam nas questões pesquisando respostas para suas dúvidas, contando com a tutoria do professor.

O trabalho transcorreu envolto em discussões e debates entre os integrantes dos grupos e o professor. Os grupos resolveram as questões em quantidades distintas: de duas a quatro questões resolvidas das seis propostas. As demais questões deveriam ser resolvidas em casa⁵.

Ao final deste quinto encontro, o professor acordou com a turma que no sexto encontro seria realizada uma avaliação individual.

O sexto encontro foi destinado à avaliação do rendimento dos alunos por meio de uma prova contendo questões de múltipla escolha, aplicada individualmente, via Moodle.

5.1.1 Direcionamentos Provenientes do Estudo Piloto

O PBL preconiza a disponibilização de um “caso clínico” acerca do que se está estudando aos grupos, que, ao analisa-lo, formulam questões norteadoras cuja pesquisa pelas respostas suscitarão o avanço do conhecimento dos próprios alunos. Entretanto, o estudo piloto mostrou que a diligência dos alunos não deve ser pressuposta pois fatores como desinteresse, apego ao modelo tradicional de ensino, resistência à mudança, inabilidade na resolução de problemas, dificuldades de trabalhar colaborativamente, excesso de disciplinas cursadas, resistência à Estatística e à informática, problemas em sumarizar e expor seus resultados de pesquisa, entre outras, podem acometer os estudantes e prejudicar o trabalho com o PBL.

Inicialmente, os Alunos não tomaram para si o protagonismo do aprendizado, esperavam que o conteúdo fosse explicado pelo professor e não lidaram bem em terem

⁵O termo “resolver em casa” significa que a atividade deve ocorrer fora da sala de aula, em contra-turno, e não necessariamente na casa do estudante.

que gerar as questões norteadoras e ainda mostraram-se excessivamente dispersos e com dificuldades de entender o processo do PBL. Além disso, o fato de terem tempo para pensarem, em grupo, sobre as questões norteadoras, parece ter gerado angústia e desinteresse em alguns alunos.

Esta resistência inicial pode estar ligada ao fato da base de dados utilizada (dados sobre diabetes em mulheres) não condizer com a realidade dos alunos. O professor escolheu esta base de dados por ser livre⁶ e por não ter encontrado outra base específica do curso de Odontologia. O sucesso do PBL está muito ligado ao interesse dos estudantes no tema proposto, e para que isto ocorra, é preciso haver uma integração entre os Professores de Estatística auxiliados pelo(s) colega(s) do curso de Odontologia. Entretanto, o contexto das atividades do estudo piloto não contou com esta integração pois os Professores do curso de Odontologia em questão não costumam compartilhar seus dados, mesmo os já publicados, desatualizados, corrompidos, incompletos, etc.

O caminho natural do PBL tende a levar bastante tempo, considerando a baixa estruturação (não há um roteiro detalhado a seguir) das atividades, porém, num contexto onde o PBL não faz parte do currículo do curso de Odontologia, este tempo pode ser demasiado.

Por este fato, não há condições de deixar que o PBL siga o seu curso dependendo exclusivamente do progresso natural dos grupos, ou seja, o papel típico apregoado ao Tutor no PBL pareceu não ser capaz de ultrapassar, por si só, as resistências observadas, senão com adequações metodológicas propostas ao PBL que precisam estar mais no sentido de estruturar o caminho de aprendizado que os estudantes irão seguir, propondo-lhes questões norteadoras prévias (para não depender somente das formuladas pelos Alunos), remodelagem e maior estruturação do conteúdo trabalhado (diferentemente do que apregoa o PBL típico).

O percurso de seis semanas para a aplicação das atividades foi grande demais,

⁶Disponível em <https://www.kaggle.com/uciml/pima-indians-diabetes-database>

suscitando adequações no PBL para a aplicação definitiva, da seguinte forma:

- Aumento da estruturação das questões norteadoras: propor mais questões norteadoras que direcionem o estudo para foco desejado, no momento da apresentação das atividades.
- Aumento do material de apoio: disponibilizar materiais em formatos distintos (escrito e áudio-visual) que desenvolvam o conteúdo dos THMP
- Aumento da estruturação do conteúdo dos TH: utilização de uma abordagem dos THMP, com as quatro etapas mostradas abaixo.
 1. Estabelecimento das hipóteses nula e alternativa;
 2. Estabelecimento do nível de significância;
 3. Cálculo do p-valor;
 4. Tomada de decisão.

Em relação ao currículo, a disciplina de Bioestatística para o curso de Odontologia prevê duas aulas teóricas, em sala de aula tradicional e duas aulas práticas, em laboratório de informática, entretanto, a decisão de utilizar somente o laboratório de informática para o desenvolvimento das atividades foi acertada, pois uniu melhor a teoria à prática e deu agilidade à pesquisa dos Alunos e às intervenções do professor/tutor.

Quanto ao instrumento para o estabelecimento do nível de apropriação do CC dos THMP, estudo piloto mostrou que este não poderia ser aplicado (como, de fato, não foi), pois lhe faltava considerar a explicitação ou não de invariantes operatórios dentro das etapas do THMP, etapas estas que ainda não estavam adequadamente ajustadas no instrumento.

Por fim, é relevante considerar que, durante a aplicação do PBL, há momentos em que o Tutor conduz o processo a partir de um questionamento feito pelo aluno. Quando

o aluno faz uma pergunta, frequentemente é útil responde-la com outra pergunta, para que o processo ativo de busca do conhecimento se preserve no aluno. Colocar a questão adequada no momento adequado respeita o nível individual de desenvolvimento de cada estudante.

Neste sentido, o questionamento feito pelo Tutor é, basicamente, reativa à manifestação do aluno de explicitação de um possível invariante operatório. A formulação da indagação de resposta pelo Tutor deve estar dentro da zona de desenvolvimento proximal subjacente ao invariante operatório explicitado. Sendo, então, relevantes para a intervenção eficiente do Tutor, além da identificação de possível invariante operatório, o diagnóstico provável da ZDP subjacente a ele e a formulação da indagação que suscite a expansão da ZDP, ou por meio da *ruptura* com o invariante operatório incorreto, ou pela *expansão* subsequente ao invariante operatório correto, observável pela análise da resposta do aluno à indagação. Todo este processo, se levado à cabo, corresponde ao que aqui chamar-se-á de Oportunidade Didática (OD).

Requer salientar que a OD inicia com o questionamento ou explicitação de alguma dúvida do aluno quando expõe um possível invariante operatório e como resposta, obtém uma indagação ou encaminhamento do Tutor que objetiva a ruptura ou expansão relativas a este invariante operatório, fechando um ciclo que pode repetir-se. Ou seja, o ciclo começa e se encerra no aluno, passando necessariamente pela ação potencializadora do Tutor, em formato de espiral ascendente.

Resumidamente, a Oportunidade Didática é a ação didática do professor (ou tutor) sobre um invariante operatório manifestado explicita ou implicitamente pelo estudante com a intenção de transforma-lo por ruptura ou reforço.

5.2 Aplicação Prévia do PBL

O estudo piloto revelou que os Alunos se ressentiram por não estarem habituados com o trabalho do PBL, principalmente no que se refere à proposição de questões norteadoras, organização do trabalho em grupo e explicitação de seus achados de pesquisa. Este fato corroborou para uma aplicação muito demorada do PBL, desta forma, no semestre subsequente, aplicou-se a metodologia PBL envolvendo o conceito de dispersão, um pré-requisito dos THMP.

Abaixo, apresenta-se um delineamento do campo conceitual do conteúdo dispersão, construído por meio da TCC. Em seguida, é relatada uma experiência didática de aplicação da PBL apoiada na TCC a uma turma da Psicologia, que deu origem ao artigo intitulado “*Problem Based Learning* para Favorecer a Apropriação do Campo Conceitual do Conceito de Dispersão por Alunos do Curso de Psicologia”.

CAMPO CONCEITUAL DO CONCEITO DE DISPERSÃO

O campo conceitual científico do conceito de dispersão (C) é constituído por um conjunto de situações (S), de invariantes operatórios (I) (estatísticos (IE) e matemáticos (IM)) e por um conjunto de representações (R), logo $C = \{S, I, R\}$.

O conjunto das situações $S = \{F \cup P\}$ inclui fenômenos (F) e problemas (P) estatísticos que descrevem e dão sentido ao conceito de dispersão relativos à área da saúde como por exemplo, a variabilidade dos escores obtidos por diferentes indivíduos em um teste psicológico.

O conjunto dos invariantes operatórios $I = \{IE \cup IM\}$, cientificamente aceitos, aplicáveis às situações, dão significado ao conceito de dispersão por meio de suas propriedades, relações e transformações. O conjunto dos invariantes operatórios estatísticos (IE) tem como elementos, as operações e propriedades que se relacionam com o conceito estatístico de dispersão como a classificação das variáveis (quantitativas), medidas de tendência central (em especial a média), amplitude total, separatrizes, desvios, desvio

médio, variância, desvio padrão, coeficiente de variação, gráfico de dispersão, boxplot, etc. que são sujeitos a princípios estatísticos como o teorema do limite central, estimadores ótimos, modelos de distribuição de frequências, etc. O conjunto dos invariantes operatórios matemáticos (IM) tem por elementos os significados matemáticos das quatro operações básicas, potência, módulo, e envolvem seus conceitos, operações, propriedades e teoremas principais.

O conjunto das representações simbólicas (R) tem por elementos os significantes estatísticos e matemáticos tanto das situações (S) quanto dos invariantes operatórios (I), de tal forma que as representações permitem externar de maneira coerente e concisa os significados relativos ao campo conceitual da dispersão. Estas representações são de caráter algébrico (quando se escreve uma fórmula ou uma expressão, por exemplo), gráfico (quando se esboça um gráfico, por exemplo), geométrico (quando se hachura a área de uma figura, por exemplo), pictórico (quando se desenha uma figura representativa da relação entre população e amostra, por exemplo) e de escrita natural envolvendo as fórmulas matemáticas, os gráficos, as distribuições de frequências, um gráfico de distribuição, etc.

Ao se levantar os dados dos escores de indivíduos em um teste psicológico, espera-se que estes escores variem em torno de um valor médio (média aritmética⁷). O espalhamento destes valores em torno da média é a dispersão dos dados e uma importante medida desta dispersão é o desvio-padrão. É possível estender esta definição usando outras medidas de tendência central em substituição da média, como a moda ou a mediana.

Cada indivíduo responde à ESV marcando uma alternativa entre sete possíveis (escala Likert de sete níveis) em cada uma das cinco questões da escala. Pode-se atribuir valores de 1 a 7 para estas opções e a soma destes valores, nas cinco questões da ESV é o escore do indivíduo nesta escala. Assim procedendo, os valores dos escores variam

⁷Será usada a palavra média para referir-se à média aritmética

entre 5 (para quem marcou a primeira opção de resposta em todas as cinco questões da escala) até 35 (para quem marcou a última opção de resposta em todas as cinco questões da escala). Desta forma, o uso do escore da ESV atende à definição de espaço amostra, qual seja, o conjunto de todos os resultados possíveis de um experimento (Díaz e López (2012);Martinez (2015)). Não é possível antever o escore de cada indivíduo, logo o experimento é dito não-determinístico e como se pode atribuir um valor numérico a cada elemento do espaço amostral, tem-se que os escores da ESV configuram uma variável aleatória unidimensional, que na notação de Díaz e López (2012) segue a definição de qualquer função, ou seja:

$$X : E \rightarrow \mathbb{R}$$

$$e \mapsto X(e) = x_e$$

Onde, X_e representa um número real a cada evento elementar, e , do espaço amostral E .

Note que, embora o escore de cada indivíduo seja um número natural entre 5 e 35 ($x \in \mathbb{R} \mid 5 \leq x \leq 35$), é conveniente tratá-lo como número real pois faz sentido pensar que um escore médio de alguns indivíduos, por exemplo, 22,5, seja melhor do que 22,8.

Considerando o escore médio de um grupo com n indivíduos (\bar{x}) como referência para o estabelecimento da dispersão, esta pode ser pensada como o distanciamento de cada escore particular (x_i) até o escore médio, chamado de desvio ($d_i = x_i - \bar{x}$). Analisar os n desvios (d_i) separadamente nem sempre é suficiente para se ter uma ideia do desvio de todo o grupo, assim, é conveniente estabelecer uma medida resumo para estes desvios.

É comum os alunos sugerirem que se faça a média dos desvios ($\frac{\sum_{i=1}^n d_i}{n}$), mas este fato tem potencialidade de ser um invariante operatório incorreto, o que indica uma OD por ruptura, onde o Tutor tem a chance de mostrar que a soma dos desvios é sempre zero, pela natureza da média aritmética, que está posicionada no centro dos dados de

forma que desvios maiores do que a média anulam-se com os desvios menores do que a média. Matematicamente, tem-se

$$\sum_{i=1}^n d_i = \sum_{i=1}^n x_i - \bar{x} = \sum_{i=1}^n x_i - \sum_{i=1}^n \bar{x} = \frac{n}{n} \sum_{i=1}^n x_i - n\bar{x} = n \frac{\sum_{i=1}^n x_i}{n} - n\bar{x} = n\bar{x} - n\bar{x} = 0$$

Assim, como a média dos desvios é sempre zero, a média destes também será sempre zero, inviabilizando o uso da média dos desvios, rompendo com a ideia inicial e intuitiva de que uma boa medida resumo dos desvios é a média dos desvios.

Neste ponto, é de se esperar que os alunos não saibam qual é a melhor medida resumo dos desvios, caracterizando nova OD, agora por expansão do entendimento dos alunos com a indagação de outras possibilidades, chegando no módulo dos desvios, o que acarreta no desvio médio ou no quadrado dos desvios, que gera a variância.

Ao usar o módulo, a ideia é deixar todos os desvios positivos (valores absolutos), permitindo o cálculo da média. Matematicamente, o desvio médio é dado da seguinte forma.

$$\frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

O uso do desvio médio, embora intuitivo, é pouco utilizado, em favor do quadrado dos desvios, que também gera uma média positiva, chamada de variância, definida por

$$\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

Existem dois tipos de cálculos para a variância, a variância populacional e a amostral. A diferença entre eles é que a variância amostral usa a ideia de grau de liberdade, que para a variância é igual a $n - 1$ pois, se algum dos n desvios for faltante (restando $n - 1$ desvios), podemos determiná-lo lançando mão da propriedade que diz que a soma dos

desvios é sempre zero e resolvendo a seguinte equação:

$$d_1^2 + d_2^2 + \dots + k + \dots + d_n^2 = 0$$

Onde k representa o desvio faltante.

Assim, as variâncias ficam definidas como abaixo:

$$\begin{aligned} \text{Populacional: } \sigma^2 &= \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \\ \text{Amostral: } s^2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \end{aligned}$$

Onde μ é a média populacional.

Por outro lado, as variâncias assim definidas, não estão na mesma unidade de medida da variável original, pois os desvios estão elevados ao quadrado. Neste caso, é conveniente extrair a raiz quadrada da variância a fim de voltar para a unidade original. O resultado desta operação é chamado de desvio padrão.

$$\begin{aligned} \text{Populacional: } \sigma &= \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \\ \text{Amostral: } s &= \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \end{aligned}$$

É comum que os alunos confundam desvio com desvio padrão, o que pode mostrar um problema de entendimento no início do processo.

Desvios maiores vão gerar desvios padrão maiores, logo, no caso dos escores da ESV, quanto maior for o desvio padrão, mais heterogêneo são os dados, ou seja, mais dispersos do centro (média) são os dados.

Por outro lado, dados que o desvio padrão está na mesma unidade da sua variável, não é correto comparar desvios padrão provenientes de variáveis distintas, ou seja, não se pode comparar o desvio padrão da idade de um grupo com o desvio padrão da altura do mesmo grupo.

Para isto, recorre-se a um coeficiente que “elimina” a unidade da variável ao comparar o desvio padrão com a média da variável, criando uma razão sem unidade e passível de comparação entre variáveis diferentes. O coeficiente de variação (CV) é dado por

$$\text{Populacional: } CV_P = \frac{\sigma}{\mu}$$

$$\text{Amostral: } CV_a = \frac{s}{\bar{x}}$$

Entretanto, a variação entre grupos, comparáveis pelo CV, não permite comparar resultados de indivíduos particulares destes grupos, como salienta [Díaz e López \(2012\)](#)

Os coeficientes de variação servem para comparar as variabilidades de dois conjuntos de valores (amostra ou populações), ao passo que se desejamos comparar dois indivíduos de cada um desses conjuntos, é necessário usar os valores padronizados. Nenhum deles possui unidade e é um erro frequente entre estudantes de Bioestatística confundi-los. ([Díaz e López, 2012](#), p. 40)

Os valores padronizados a que se refere o autor é uma comparação da distância de um determinado dado até a média dos dados com o desvio padrão destes dados, comumente chamado de escore Z:

$$\text{Populacional: } Z_P = \frac{x_i - \mu}{\sigma}$$

$$\text{Amostral: } Z_a = \frac{x_i - \bar{x}}{s}$$

Notadamente, o escore Z mede a distância do dado até a média, em termos de desvios padrão.

Pela forma que o escore Z é definido, sua média assume valor zero enquanto que o seu desvio padrão assume valor um.

Porém, [Urbina \(2009\)](#) lembra que é comum em testes psicológicos, utilizar-se escores com outra média e desvio padrão, por meio da fórmula:

$$\text{Novo escore padrão} = (\text{novo desv. pad.})(\text{escore Z}) + \text{nova média}$$

Um escore que utiliza esta estrutura é o escore T, que segundo Cabral (1996), é um escore derivado, expresso numa escala com média 50 e desvio padrão 10, dado por $T = 10 \cdot \frac{x-\mu}{s+50}$, onde x é o escore bruto, μ é a média e s é o desvio padrão da distribuição dos escores brutos.

Estando o campo conceitual do conteúdo de dispersão bem definido, passa-se a relatar uma experiência didática proposta a alunos de Psicologia.

Contexto de Ensino

A ação pedagógica foi desenvolvida com os 58 alunos participantes do PBL, porém, escolheu-se, por sorteio, a turma PSC-A para ser analisada nesta seção, visto que este desenvolvimento não é foco principal desta Tese, ficando a análise de todas as turmas, extensa demais.

A turma PSC-A tinha 25 alunos (15 mulheres e 10 homens), na disciplina de bioestatística, ofertada no terceiro semestre do curso de Psicologia de uma universidade do Rio Grande do Sul, Brasil.

Os momentos presenciais ocorreram durante três encontros semanais com três períodos de 50min cada encontro, em laboratório de informática, com acesso livre à internet, planilha eletrônica EXCEL v. 2016 e *software* estatístico Bioestat⁸.

Ao chegarem neste ponto de aplicação das atividades, os alunos tinham familiaridade com planilhas de dados no Excel, classificação de variáveis, tabelas de frequências, gráficos de colunas e histograma e medidas de tendência central, quartis, percentis, gráficos, etc, que também fazem parte do CC da dispersão e o percentil, particularmente, é utilizado na análise de ESV e costuma gerar dúvidas nos estudantes que frequentemente não levam em conta que os dados precisam estar ordenados e que o percentil divide os dados em cem partes iguais, assim, o percentil 72, separa os 72% dos menores escores dos 28% dos maiores escores, no caso da ESV.

Os alunos foram organizados em grupos de 5 integrantes (ao todo foram cinco

⁸Disponível em: <https://www.mamiraua.org.br/downloads/programas/>

grupos, referidos por: A, B, C, D e E), sendo um integrante o coordenador do grupo, responsável por manter o foco e organizar o trabalho, entregar as produções escritas do grupo pelo Moodle, tomar decisões quando o grupo não entra em consenso, etc.

A situação estudada durante o PBL, foi inspirada no instrumento chamado Escala de Satisfação de Vida (ESV) dos autores Claudio Simon Hurtz, Cristian Zanon e Marucia Patta Bardagi, do Laboratório de Mensuração da UFRGS, apresentado no livro Avaliação em Psicologia Positiva organizado por Claudio Simon Hutz, editora Artmed, 2014.

O tema da situação geradora do PBL foi escolhido em conjunto com uma Professora do curso de Psicologia da mesma IES e que tinha conhecimento da turma onde seria aplicada a metodologia. A Professora também era responsável pelo Laboratório de Estudos e Avaliação Psicológica (LEAP), do curso de Psicologia.

A ESV é composta por cinco questões de múltipla escolha em escala do tipo Likert de 7 níveis, onde o nível 1 corresponde a “discordo plenamente” e o 7 a “concordo plenamente”. Para a análise da resposta do indivíduo calcula-se o escore total de suas respostas e compara-se com uma tabela contendo escores brutos, percentis e escores t de uma amostra de referência.

As instruções para aplicação e as questões da escala são apresentadas na figura 5.2, e o texto completo que a embasa está no apêndice E.

Procedimentos Didáticos

As questões da ESV (5.2) foram disponibilizadas no Moodle para que os próprios alunos as respondessem antes da aula presencial, formando a base de dados para o PBL.

Na aula subsequente, o professor/tutor disponibilizou o capítulo do livro que embasa a ESV (E), acrescida de algumas questões norteadoras do estudo (inseridas como comentário no texto em pdf), objetivando promover um debate em sala de aula e auxiliar na interpretação da escala, são elas:

ESCALA DE SATISFAÇÃO DE VIDA
Laboratório de Mensuração da UFRGS **Instruções**

Abaixo você encontrará cinco afirmativas. Assinale na escala abaixo de cada afirmativa o quanto ela descreve a sua situação pessoal. Não há respostas certas ou erradas, mas é importante você marcar com sinceridade como você se sente com relação a cada uma dessas afirmativas.

1) A minha vida está próxima do meu ideal.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

2) Minhas condições de vida são excelentes.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

3) Eu estou satisfeito com a minha vida.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

4) Até agora eu tenho conseguido as coisas importantes que eu quero na vida.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

5) Se eu pudesse viver a minha vida de novo eu não mudaria quase nada.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

Figura 5.2: Escala de Satisfação de Vida.

- Se é subjetiva (a ESV), como pode ser medida com números, que são bastante objetivos?
- Mesmo que uma pessoa seja o mais sincera possível ao responder a ESV, isto significa que sua resposta reflete a realidade?
- Esta (a ESV) é uma escala do tipo Likert com 7 níveis. Quais as ferramentas estatísticas você usaria para analisar escalas deste tipo?

O texto que embasa a ESV (E) informa uma tabela com dados de uma amostra de referência para análise da satisfação de vida, mas não explica o que são alguns destes dados, assim, além das questões acima, perguntas menos subjetivas e referentes ao conteúdo estatístico envolvido no texto da ESV foram propostas para serem analisadas

em grupo:

- O que é um percentil?
- Quais são os percentis, na opinião do grupo, que encerram os limites baixo, médio e alto de satisfação de vida?
- O que é escore-t?
- Qual é o significado destas letras e de seus valores ($M=21,8$ e $DP=7,3$)?

Estas questões, estruturaram o estudo da ESV nos grupos, preparando-os para o segundo encontro.

No segundo encontro, os grupos já tinham feito suas pesquisas acerca das questões norteadoras e puderam tirar dúvidas com o professor/tutor durante o primeiro momento do encontro.

Em seguida, o professor/tutor apresentou pergunta que seria o objetivo da análise estatística que os grupos iriam realizar:

De acordo com os dados provenientes da ESV, como está a satisfação de vida dos alunos da turma?

A análise estatística deveria ser entregue de forma escrita (em Word) e conter o que lhes fosse importante para responder às questões. Não houve outro tipo de exigência.

Para estruturar melhor a pesquisa dos alunos foram disponibilizados materiais didáticos de apoio como textos, planilhas, vídeos, enquetes e avaliações, por meio do ambiente virtual de aprendizagem Moodle.

Apresentação e Discussão dos Resultados

Os dados qualitativos que serão apresentados e discutidos abaixo, foram levantados por meio do diário de classe do professor/tutor e de um estudo avaliativo das análises estatísticas produzidas pelos Alunos.

O problema escolhido para o PBL pareceu adequado pois os alunos, em geral, mostraram motivação em aprender sobre um instrumento utilizado em sua profissão. Porém, o processo do PBL não foi entendido pelos alunos, que em geral, pareciam estar lidando com um trabalho em grupo tradicional, que já estavam acostumados a fazer.

No primeiro encontro os alunos trabalharam bem na atividade, perguntando e participando ativamente das discussões. No início as perguntas foram mais do tipo o que deviam fazer, mas com o tempo, a criatividade dos grupos foi aparecendo e as perguntas passaram a ser se o que fizeram estava certo. Ao professor/tutor, coube orientar, por meio de novas perguntas, a fim de que o próprio grupo justifique suas respostas para além do certo ou errado e, assim, entrassem de forma mais autônoma no processo do PBL.

Relativamente às análises estatísticas, os grupos utilizaram diversos recursos estatísticos relacionados à dispersão. Um destes recursos é o gráfico de dispersão, bastante explorado pelos grupos. O grupo A apresentou dois diagramas de dispersão, mostrados na figura 5.3.

Por sua vez, o grupo GB trouxe um tipo diferente de diagrama de dispersão onde os dados foram ordenados e uma linha com a média da amostra foi introduzida, como se vê na figura 5.4, mas não julga se os valores estão muito o pouco dispersos. Os grupos D e E também ordenaram os dados no diagrama de dispersão mas não colocaram uma linha com a média e o grupo C inseriu um diagrama de dispersão sem ordenação.

O grupo D interpretou a dispersão por meio do diagrama (figura 5.5), identificando uma não tão grande dispersão pela baixa frequência dos valores de escore mínimo e máximo, apresentados no gráfico. Esta interpretação parece indicar a presença do teorema em ação “amplitude total baixa implica em baixa dispersão”, que é incorreto por estar incompleto.

Muitas tabelas de frequências também foram utilizadas nas análises, principalmente para apontar o acúmulo de indivíduos em uma determinada faixa de escores, como fez

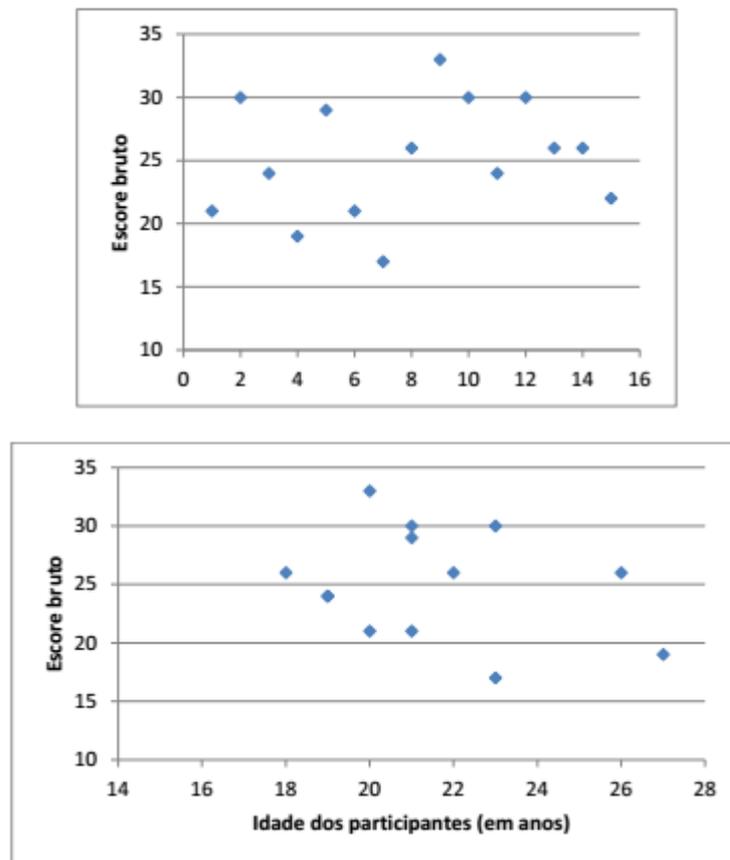


Figura 5.3: Dois gráficos de dispersão apresentados pelo grupo A.

o grupo B (figura 5.6).

O grupo D fez uma tentativa de expansão da interpretação do histograma (figura 5.7) mas confundiu a classe modal com a moda, explicitando um possível conceito em ação: “a classe modal é a moda da distribuição”, que está incorreto pois moda é o valor (ou os valores) com maior frequência e não a classe com maior frequência da distribuição. O grupo D também afirma, temerariamente, que a classe modal encerra escores entre 4 e 6 na escala Likert da ESV, sem apontar onde está a informação, fazendo parecer que se referiam ao histograma em questão, que, por si só, não permite esta conclusão.

Os alunos mostraram que conseguem analisar características da dispersão por meio de um *boxplot*, fazendo supor-se que existe um conceito em ação que relaciona *boxplots*

O **diagrama de dispersão** mostra em pontos cada valor, para vermos como estão dispersos. A linha vermelha é a média (25,2). Podemos ver que o número de valores acima da média são 8, e os abaixo da média são 7.

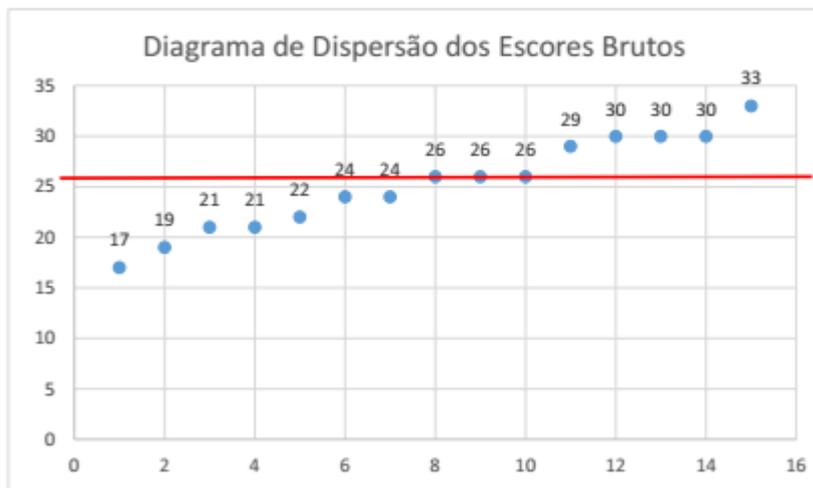


Figura 5.4: Gráfico de dispersão com explicação do grupo B.

a análise da dispersão. O grupo C analisou a dispersão por meio do *boxplot*, como se vê na figura 5.8.

A análise do *boxplot* feita pelo grupo D pode ser vista na fig 5.9.

A explicação dos alunos parece encerrar o seguinte teorema em ação: se o retângulo do *boxplot* for simétrico e com pouca amplitude então os dados têm baixa dispersão. Este teorema em ação está parcialmente correto visto que a simetria em relação à mediana em um *boxplot* não implica em baixa dispersão, porém, a proximidade entre os quartis um e três (chamada de pouca amplitude, pelos alunos) impacta na baixa dispersão dos dados.

As medidas de dispersão, que poderiam corroborar com a análise dos diagramas de dispersão, também apareceram, prioritariamente em seu caráter procedimental (mecânico) e teórico e timidamente em seu sentido avaliativo (interpretativo) exceto pelas manifestações dos grupos B e D, mostrados na figura 5.10.

O score-T, além de importante para a ESV, envolve o entendimento do conteúdo

Avaliando o Diagrama de Dispersão (GRÁFICO 3), colocado com os Escores Brutos em ordem crescente, é possível observar a variação das respostas (escores brutos), que vão de 17-33. Assim, apesar de existir uma variação nas respostas, ela não é tão destoante, já que foram apenas duas pessoas que somaram os escores (uma o mínimo e uma o máximo).

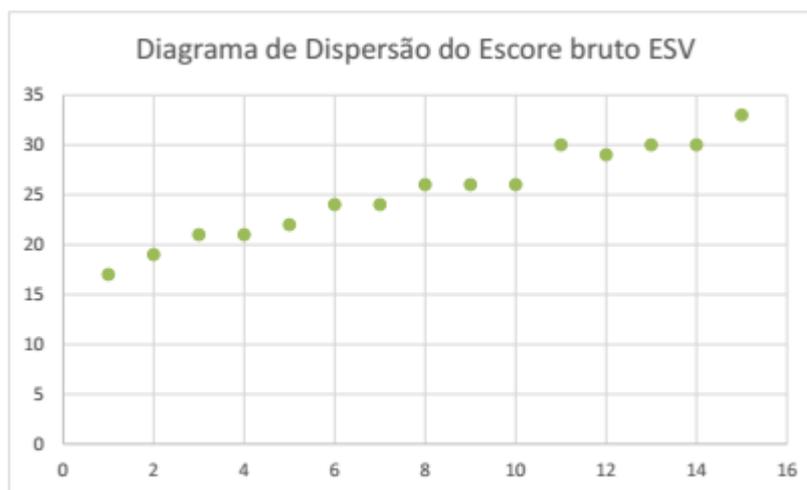


GRÁFICO 3: Diagrama de dispersão das respostas

Figura 5.5: Interpretação e gráfico de dispersão do grupo D.

de dispersão, entretanto, os alunos não relacionaram o escore-T com o desvio padrão e a média de maneira completa preferindo abordá-lo em sua forma operatória ou teórica. Os grupos A e B apresentaram a fórmula que relaciona o escore Z com o T ($T = Z \cdot 10 + 50$), calcularam os escores T da amostra mas não interpretaram os resultados.

Chama atenção que, mesmo envolto num ambiente de pesquisa, haja manifestações, como a do grupo A, cuja análise estatística trouxe uma série de cálculos e gráficos, inclusive envolvendo fórmulas que não foram trabalhadas em aula, mas sem nenhum tipo de análise.

Por outro lado, a análise do grupo D contrastou com as demais ao apresentar uma coerência entre os resultados, os conceitos e as conclusões. As conclusões foram embasadas pelos conceitos estatísticos aplicados sobre os dados, levando em conta mais de um conceito acerca da dispersão, como se vê no seguinte excerto da figura 5.11.

Histograma: Divide em grupos, os resultados, vemos que de 23 até 26 existem 5 escores brutos, por exemplo.

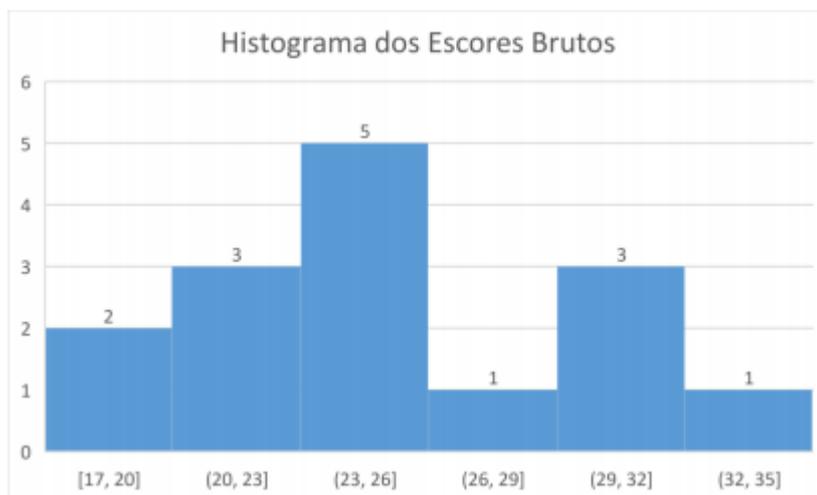


Figura 5.6: Interpretação do histograma dos escores do grupo B.

Desta forma, o grupo D mostrou uma competência na análise da dispersão que junta várias habilidades, neste caso, o uso do desvio padrão, do coeficiente de variação, do diagrama de dispersão e da amplitude. O que remete à existência de conceitos em ação relevantes para o conceito mais abrangente da dispersão.

Cabe ressaltar que invariantes operatórios dizem respeito ao indivíduo, e não ao grupo, entretanto, as ideias do grupo explicitadas no trabalho, provêm dos seus indivíduos, portanto, o invariante operatório em questão pode pertencer, completa ou parcialmente, a mais de um integrante.

Tendo em vista as análises estatísticas produzidas pelos Alunos, transparece a ideia de que o conceito de dispersão está presente, forte porém implicitamente, nos Alunos que o expõem prioritariamente pelo seu caráter procedimental, por meio do cálculo do desvio padrão, escore Z, escore T, amplitude total e, mais raramente, coeficiente de variação e percentis. Os Alunos também usaram os gráficos de dispersão e histogramas como forma de explicitação do conceito de dispersão. Assim, a rede de conceitos que

Avaliando o histograma (GRÁFICO 1), nota-se que 5 dos avaliados têm o escore bruto entre 23-26 (que é a Mo, ou seja, os escores que mais se repetiram), que é um dado relativamente positivo, já que as respostas do ESV destas pessoas estão entre 4-6 (aproximadamente).

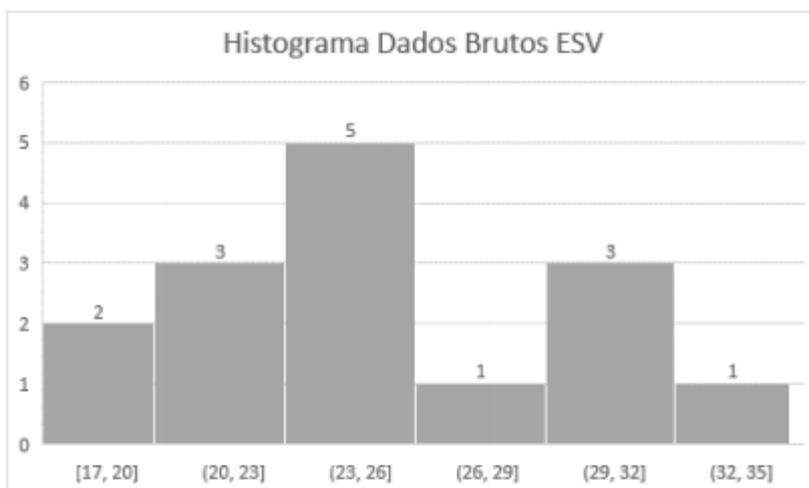


GRÁFICO 1: Histograma dos dados brutos da ESV.

Figura 5.7: Interpretação do histograma dos escores do grupo D.

dão sentido ao conceito de dispersão parece estar ampla o suficiente dentro deste CC, muito embora, julgamentos do tipo “a dispersão é grande” ou “a dispersão é pequena”, tenham sido evitados pelos alunos.

Concluída esta etapa prévia à aplicação definitiva das atividades do PBL acerca dos THMP (capítulo 6), julga-se satisfatória a preparação dos estudantes para a etapa seguinte, fonte dos dados desta pesquisa.

Boxplot: é um tipo de gráfico usado para avaliar a distribuição de dados, é formado pelo primeiro e terceiro quartil e a mediana. Ele fornece informações importantes sobre a variabilidade dos dados. Nosso gráfico boxplot ficou quase simétrico, tendo pouca amplitude, já que a linha da mediana está próxima do centro do retângulo. Logo, os dados não variam muito.

Figura 5.8: Interpretação do *boxplot* dos escores do grupo C.

Ao analisar o Boxplot (GRÁFICO 2), podemos concluir que ele é proporcional e simétrico, o que significa que não há números que destoam muito dentre os avaliados. A mediana é 26, ou seja, é o número que está no centro dos valores e é aproximada da média, ratificando a concordância do grupo. A média das respostas é de $\bar{x}=25,2$, o que significa um bom índice de ESV na turma.

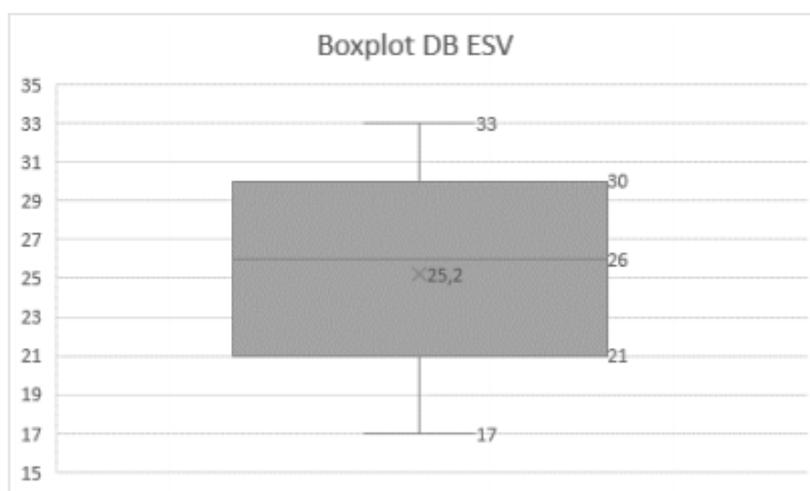


GRÁFICO 2: Boxplot dos dados brutos da ESV do grupo analisado.

Figura 5.9: Interpretação do *boxplot* dos escores do grupo D.

Desvio Padrão: podemos ver pelo desvio padrão dos escores brutos, podemos ver se os índices estão muito dispersos, ou longe da média. Quanto mais próximo de 0 o desvio padrão, mais os valores são uniformes. Como o desvio padrão resultou 4,47, os índices são pouco uniformes ou heterogêneos.

O Desvio Padrão é de 4.47 pontos na ESV, o que é um DP baixo, e o Coeficiente de Variação é de 18%. Analisando estes dados, podemos concluir que o grupo é homogêneo, e que a variabilidade, embora exista, é baixa.

Figura 5.10: Definições de desvio padrão (acima grupo B, abaixo grupo D).

É possível perceber que os indivíduos avaliados na turma possuem o índice de satisfação de vida elevado comparada à Escala Padrão para Adultos, já que ninguém apresentou ESV abaixo do percentil 25, tendo como escore bruto mínimo 17 (na escala padrão o EB mínimo é de 9). Também teve um número positivo de pessoas com o escore 30, que é um escore elevado. Além disso, a média do grupo aqui analisado ($\bar{x} = 25.2$) é maior que a do grupo padrão ($\bar{x} = 21.8$). Isso pode ser pensado como possível indicador do nível social ocupado pelos estudantes universitários de uma universidade particular, possivelmente mais alto que do grupo analisado, o que também faz com que o DP deste seja menor, indicando maior homogeneidade do grupo.

Figura 5.11: Interpretação da variação dos escores da ESV do grupo D.

Capítulo 6

Aplicação das Atividades: A Metodologia de Ensino

Neste capítulo, descreve-se a metodologia de ensino que foi adotada, considerando as mudanças e adaptações propostas desde o estudo piloto. A análise dos resultados desta etapa está na seção 7.1. Durante a descrição da metodologia, recorre-se às fundamentações teóricas que embasaram sua construção, expondo algumas delas.

Cabe lembrar que o PBL apoiado pela TCC, tal como será descrito adiante, foi aplicado a uma turma do curso de Nutrição com dez alunos (NUT-A) e a duas turmas do curso de Psicologia (PSC-A e PSC-C) com 26 e 25 alunos, respectivamente. Esta aplicação foi feita na sequência cronológica da aplicação prévia (seção 5.2) e foram mantidos os grupos de alunos estabelecidos naquela oportunidade. Os grupos são importantes no PBL. Duch et al. (2001) aconselha que o trabalho em grupos inicie cedo no curso, o que foi feito.

Desta forma, havia familiaridade com a metodologia de ensino e os conhecimentos prévios aos THMP já haviam sido trabalhados, como mostrado na seção 5.2. Todos os encontros ocorreram em laboratório de informática.

O estudo piloto (seção 5.1) mostrou a necessidade de um aumento na estruturação

das atividades do PBL na tentativa de se otimizar o tempo de aplicação, muito extenso para a realidade da IES em questão. Neste sentido, o professor/tutor estabeleceu questões norteadoras iniciais e organizou materiais de apoio sobre os THMP, tal como já havia feito na aplicação prévia com o conteúdo de Dispersão, e os disponibilizou via Moodle, como se vê nos anexos [F](#) e [G](#).

Aqui, escolheu-se apresentar a metodologia de ensino considerando a aplicação nas turmas do curso de Psicologia, bem maiores do que a turma do curso de Nutrição, com o objetivo de não tornar a leitura enfadonha, visto que a proposta é a mesma, mudando somente o campo de aplicação das situações.

Desta forma, a metodologia adotada para o ensino dos THMP foi a seguinte:

Primeiro Encontro O professor/tutor apresentou a situação principal (comumente chamada de caso clínico, no PBL) e as primeiras questões norteadoras, bem como o material de apoio e lembrou as vantagens do trabalho em grupo, como [Duch et al. \(2001\)](#) sugere fazer no início das atividades, ressaltando que o grupo é a equipe formada para a resolução dos problemas decorrentes das situações e que sua força está no trabalho colaborativo, responsável e empático, tal como coloca [Yew e Goh \(2016, p. 76\)](#), “a filosofia subjacente do PBL é que o aprendizado pode ser considerado uma atividade “construtiva, autodirigida, colaborativa e contextual”.

Os materiais ficaram disponíveis aos alunos no Moodle e podem ser vistos nos anexos [F](#) e [G](#).

O professor/tutor lembrou aos alunos que as atividades aconteceriam em sala de aula mas se estenderiam para fora dela como atividades extra-classe e que não se restringiriam às respostas às questões iniciais já propostas, senão ao envolvimento com a situação principal, em cujo processo, os THMP seriam compreendidos, de forma ativa, por eles, estudantes. Deste modo, as questões iniciais serviriam

como exemplos de novas questões a serem propostas pelos grupos, divididas entre seus integrantes que compartilhariam seus achados com o grupo no decorrer do processo.

Neste encontro, os grupos trabalharam no entendimento da situação principal e das quatro questões iniciais, reproduzidas aqui tal como estão colocadas no anexo [F](#):

Situação principal: O Professor de Estatística quer avaliar o nível de Satisfação de Vida dos alunos de Estatística de sua Universidade. Para isto aplicou a Escala de Satisfação de Vida (ESV) para os alunos que se voluntariaram, o que gerou a amostra cujos dados estão no arquivo “ESV todos os alunos com escore”.

A situação principal foi criada levando-se em conta o fato de que os estudantes já tinham um contato prévio com a ESV, inclusive colaborando com suas próprias respostas à Escala. Eles já haviam estudado este conjunto de dados quanto à estatística descritiva, na aplicação prévia. Desta maneira, leva-se em conta a zona de desenvolvimento proximal do aluno ao propor-lhes uma atividade investigativa, tal como preconiza o PBL, propensa a movimentar seu esquemas de pensamento no sentido da evolução, por reforço ou ruptura, de seus invariantes operatórios estabelecidos até aqui.

Com esta situação aberta, afasta-se da metodologia tradicional de ensino por não se esperar que os alunos encontrem uma resposta previamente conhecida, o professor não expõe o conteúdo antes de propor a situação e os alunos precisam pesquisar para além das questões iniciais e colaborarem entre si para aprenderem sobre THMP.

Neste sentido, o ambiente investigativo, preconizado pelo [ASA Revision Committee \(2016, p. 3\)](#), reforça a assertiva de que os alunos aprendem pela construção do conhecimento, feita por [Garfield e Ben-Zvi \(2007\)](#), além de considerar o campo

de atuação dos alunos, exigência do PBL. Desta maneira, ao se propor esta situação, respeita-se a ZDP dos estudantes, que será o campo onde os invariantes operatórios estarão propensos a se externalizarem e, ao fazê-lo, permitem que o professor/tutor lance mão de Oportunidades Didáticas catalizadoras de rupturas ou acomodações típicas do desenvolvimento cognitivo rumo ao melhor domínio do CC dos THMP.

Ao final do primeiro encontro, o professor/tutor reforçou a necessidade de que os integrantes dos grupos entrem em contato no contra-turno¹ para que as pesquisas individuais não fiquem isoladas do contexto do grupo. Esta comunicação extra-classe não foi monitorada, mas seus efeitos foram verificados nos encontros presenciais e registrados na seção 7.1.

Além da situação principal, foram disponibilizadas algumas questões iniciais, propostas pelo professor/tutor. As questões e breves comentários sobre seus objetivos didáticos estão abaixo:

1. Qual é a média da amostra?

Esta questão tem a vantagem de lincar o trabalho que os alunos estão iniciando com o que foi feito anteriormente sobre o conteúdo de Dispersão, descrito na seção 5.2. Esta familiaridade representa um ponto de partida simples, visto que os estudantes demonstraram facilidade com o cálculo e entendimento do conceito de média, estando este conteúdo, dentro da ZDP dos alunos a esta altura.

2. Considerando que um escore 25 na ESV indica que se está satisfeito com a vida. Existe alguma chance (probabilidade) da população ter média 25, mesmo que a amostra não tenha média 25? Esta chance é grande ou pequena?

Esta tarefa visa promover a discussão a respeito da diferença de uma amostra

¹Termo usado para trabalho fora da sala de aula, em “casa”.

e a população da qual foi extraída. Esperava-se que os alunos relacionassem estes conceitos com o de Distribuição Amostral da Média (DAM) trabalhado anteriormente, imediatamente antes destas atividades. Aqui, não se esperava que os alunos partissem para o cálculo de probabilidades, senão, ponderassem sobre as diversas possibilidades de se extrair uma amostra com média diferente de 25 de uma população com média 25, ponderação importante para os THMP.

3. Se a população tiver média 25, existe alguma chance de se retirar uma amostra cuja média não é 25?

Para esta questão esperava-se que os alunos continuassem a explorar as características da amostragem por meio da linguagem natural, fazendo inferências informais e ponderações livres de processos roteirizados ou pré-estabelecidos.

4. Se a população tiver média 25, qual a chance de tirar uma amostra com a média que obtivemos? Esta chance é grande ou pequena? O que significa grande ou pequeno neste contexto?

A questão continua trabalhando com os conceitos da questão anterior, aproximando-se mais das respostas que os THMP podem dar por meio do p-valor bem como o estabelecimento de um valor limite para o que se julga grande ou pequeno neste contexto, aludindo ao conceito de nível de significância, que virá mais adiante.

As quatro questões iniciais mostradas até aqui foram trabalhadas durante o primeiro encontro e para o trabalho extra-classe os alunos tinham à disposição as demais questões norteadoras:

5. Conduza um teste de hipóteses para uma média com base na amostra levantada e decida se a média populacional é de 25.
 - (a) Estabeleça a hipótese a ser testada e a alternativa. Escreva, com suas

palavras, o que elas significam dentro do contexto desta situação.

- (b) Defina o nível de significância, justifique a sua escolha e explique o seu significado dentro do contexto da questão.
- (c) Defina a região de aceitação e rejeição de H_0 .
- (d) Calcule a estatística do teste e a localize no gráfico desenhado em c).
- (e) Calcule o valor de p adequado e explique o seu significado dentro do contexto da questão.
- (f) Decida sobre rejeitar ou não a hipótese nula e estruture uma resposta escrita para a situação.

Neste ponto, as questões assumem a notação e estrutura próprias dos THMP, na busca da ligação dos conceitos trabalhados nas primeiras quatro questões com a estrutura que se quer dominar dos THMP. Adicionalmente, esta nomenclatura facilita a pesquisa dos grupos durante o trabalho em contra-turno, bem como a divisão do trabalho, típico do PBL.

Embora não se esperasse que os alunos relacionassem explicitamente as primeiras questões com estas, as ideias que foram trabalhadas anteriormente poderiam se preservar, em alguma medida, neste momento de maior rigor conceitual e estrutural.

O formato aqui exposto mantém o caráter investigativo preservado, estimula o trabalho na fronteira da ZDP dos estudantes e incentiva a exposição de invariantes operatórios entre os membros do grupo. Neste sentido, o PBL continua alinhado com os preceitos de um bom trabalho de ensino de estatística, colhidos na revisão de literatura, com a vantagem de estar apoiado pela TCC.

No trabalho em contra-turno, os alunos tinham a seu dispor, para pesquisa, os materiais de apoio: uma série de hipertextos e de vídeo-aulas coletadas da internet

ou criadas pelo professor/tutor, que apresentam alguns conceitos e exemplos do CC dos THMP, sem, porém, dar as respostas das perguntas norteadoras, forçando que houvesse pesquisas complementares e um esforço compilatório do conteúdo, por parte do aluno e dos grupos, corroborando com [Duch et al. \(2001, p. 59\)](#), quando afirmam que “No *Problem Based Learning*, os estudantes são chamados a trabalharem juntos para analisar e resolver problemas e comunicar, avaliar e integrar informação de diversas fontes”.

O material de apoio levou em conta as sete categorias definidas em consequência dos resultados apontados pelo estudo piloto: Hipótese Nula, Hipótese Alternativa, Nível de Significância, escolha da Estatística do Teste, cálculo da Estatística do Teste ou p-valor, Decisão e Interpretação.

Resumidamente, no primeiro encontro os alunos receberam a situação principal, as primeiras questões norteadoras e o material de apoio. Foram trabalhadas as primeiras quatro questões norteadoras em sala de aula e as demais precisaram ser trabalhadas em contra-turno, preparando o segundo encontro.

Sendo assim, a metodologia exposta aqui, cumpre com o que se preconiza a seguir.

[...] constata-se que o uso de metodologias baseadas em teorias de aprendizagem, desenvolvidas em ambientes que valorizam a investigação sobre dados reais e amparadas por recursos computacionais desenvolvidos em ambientes híbridos (presencial e virtual), parece estar de acordo com a evolução que o ensino de Estatística demanda e que pode melhorar o aprendizado de estudantes da área da saúde ([Fioravanti et al., 2019, p. 92](#)).

A este ponto, os alunos dispunham de todo o material disponibilizado pelo professor/tutor e, a partir daqui, a metodologia de ensino focou mais no caráter orientador do professor/tutor, como descrito a seguir.

Segundo Encontro A primeira parte deste encontro foi dedicada ao compartilha-

mento dos resultados das pesquisas individuais que os alunos realizaram em contra-turno. Um tempo bastante grande foi dedicado a este fim, com grande demanda dos grupos pelo professor/tutor, durante o processo, que os orientou no sentido de organizar suas pesquisas, evitando dar respostas prontas e procurando esta atento às Oportunidades Didáticas.

Novamente, este papel norteador e organizador do trabalho, assumido pelo professor/tutor, é uma demanda do PBL que afasta-se do modelo tradicional de ensino, que foca a ação do professor na explicação/elucidação da dúvida do aluno, dando-lhe as respostas pedidas.

Na segunda metade deste encontro, o professor/tutor disponibilizou aos grupos outras situações envolvendo os THMP, visíveis no anexo H. Com estas questões adicionais, adaptadas de livros textos de Bioestatística, facilita-se o entendimento dos THMP durante as buscas dos alunos por este conteúdo, visto que houve grupos que preferiram trabalhar primeiro as questões extras e depois retomarem a situação principal.

A primeira situação apresenta o texto da Escala de Autoeficácia Geral (EAG), que contém algumas inversões na escala likert de suas respostas, comuns neste tipo de escala, mas que ainda não haviam sido abordadas. A interpretação da EAG é semelhante à ESV, com escores totais, escore t e percentis. Os Alunos utilizaram um banco de dados fornecido em Excel para fazerem a análise da EAG. Esta questão traz um ambiente semelhante à situação principal, mas abordando alguns conceitos de forma mais diligente, como o conceito de percentil, por exemplo.

As demais questões referem-se a aspectos mais mecânicos da resolução de problemas deste tipo, como se vê a seguir.

A segunda situação traz um valor populacional de referência e pede para testar se os indivíduos de uma amostra desta população têm valores coerentes com o

valor de referência, a 1% de significância.

A terceira situação se quer testar a afirmação da OMS sobre a taxa média de suicídios para cada 100 mil pessoas com base em uma amostra com 24 elementos. O nível de significância não é pré estabelecido.

A situação número quatro procura testar se um escore populacional é menor do que um dado valor de referência baseado em uma amostra pequena e com nível de significância de 10%.

Por fim, a situação número cinco apresenta os dados amostrais de uma amostra pequena para testar se são provenientes de uma população com determinada média de referência, a um nível de significância de 5%. A situação não possui um contexto específico e apresenta a hipótese nula e alternativa. Esta situação é a mais genérica da lista.

As cinco situações são abordagens distintas, porém típicas, dos THMP e foram aplicadas para aumentar a variabilidade de situações que dão sentido ao conceito, além de todas aquelas situações que os grupos entraram em contato, por suas próprias pesquisas, durante o PBL. Desta forma, a estruturação do PBL é novamente feita com o amparo da TCC, fortificando uma possível fragilidade da PBL neste sentido, pois, se as situações dependerem somente das pesquisas dos alunos, não sejam suficientes para a mobilização dos invariantes operatórios nem dos conceitos científicos contidos no CC dos THMP.

Garfield e Ben-Zvi (2007) apud Fioravanti et al. (2019, p. 75) já apregoavam que “Estudantes aprendem melhor se recebem feedbacks consistentes que auxiliem em sua performance”. Este é um ponto importante para o PBL, na figura do tutor, e foi explorado durante os encontros presenciais por meio da busca por Oportunidades Didáticas, entretanto, os alunos não procuraram o professor/tutor durante os trabalhos em contra-turno, mesmo a plataforma Moodle permitindo

esta interação, o que dificultou a orientação do professor/tutor aos alunos, no contra-turno, sobrecarregando os encontros presenciais.

Cabe ressaltar que o objetivo desta Pesquisa esteve no entendimento do nível de domínio do campo conceitual (CC) dos THMP dos Alunos amostrados e na comparação de seus rendimentos com Estudantes que não participaram do PBL. O estudo do processo de domínio deste CC pertence aos prolongamentos desta Pesquisa.

Terceiro Encontro No terceiro encontro, os grupos apresentaram suas respostas à situação principal de acordo com a dinâmica: um grupo era sorteado e apresentava a sua resposta, em seguida, outro grupo era sorteado e apresentava as diferenças de suas respostas em relação ao primeiro grupo. O processo seguiu desta forma até cobrir todos os grupos, com a apresentação do primeiro grupo demorando mais do que as demais, como o esperado.

Ao final do terceiro encontro, ficou combinado com a turma que o quarto encontro terminaria com uma avaliação individual, contendo duas questões abertas sobre THMP, com consulta livre aos materiais individuais (anotações particulares), internet (navegadores) e aos *softwares* Excel e Bioestat, ficando proibidos o uso do telefone celular e softwares de mensagens. As questões foram aplicadas de forma escrita e geraram os dados cuja descrição e análise estão na seção 7.3.

Quarto Encontro O quarto encontro começou com o tratamento das dúvidas remanescentes do encontro anterior. O professor/tutor visitou cada um dos grupos para auxiliar no esclarecimento das dúvidas.

O encontro terminou com uma avaliação contendo duas questões, já analisadas na subseção 4.2.3 e mostradas a seguir.

1. *A taxa de fenilalanina no soro, em indivíduos normais, é de 1,4 mg/100mL*

e o desvio padrão é de 0,32 mg/100mL. Supõe-se que cem pacientes formam uma amostra aleatória da população de indivíduos normais, com média amostral de 1,3 mg/100mL.

Considerando um nível de significância de 1%, os dados estão de acordo com a suposição feita?

- 2. A média de acidentes mortais em uma cidade é de 12 mensais. Após uma campanha de conscientização sobre os perigos no trânsito, nos seis meses consecutivos contabilizaram-se 8, 11, 9, 7, 10 e 9 acidentes mortais. A campanha foi efetiva?*

Quinto Encontro O quinto encontro foi destinado à avaliação constituída de dez questões de múltipla escolha aplicadas pelo Moodle, de forma individual e com as mesmas restrições da avaliação anterior. Esta avaliação gerou os dados para a análise do rendimento dos alunos, tal como descrita em 7.2. As questões do instrumento podem ser vistas no anexo I.

Cabe lembrar que o processo do PBL não preconiza a aplicação sistemática de atividades por parte do professor/tutor, em vez disso, as pesquisas dos alunos encaminham o processo progressivo de domínio do CC em questão. Assim, o primeiro encontro ganha uma importância maior em relação aos demais, pois é ali que a questão principal é colocada e as orientações gerais são dadas, como mostrado acima. Por este motivo, a descrição dos primeiros encontros ganhou mais detalhes do que os demais.

Neste sentido, a estruturação proposta nesta pesquisa teve-se em dois focos: uma organização prévia de alguns materiais de pesquisa aos alunos e no formato do conteúdo dos THMP, ambas estruturações com o objetivo principal de permitir a aplicação do PBL num contexto de ensino tradicional e por um período de tempo bastante limitado.

A estruturação dos materiais de pesquisa, por si só, não permitia o entendimento completo do CC dos THMP mas prevenia a fuga do tema no decorrer das pesquisas.

Por sua vez, a estruturação do conteúdo procurou pôr em linhas mais amigáveis um conteúdo extenso demais para ser aprendido prioritariamente pelas pesquisas dos alunos, mesmo orientados por um professor/tutor.

Neste ponto, após ter-se pormenorizado a metodologia de ensino adotada, e apontado algumas diferenças com a metodologia tradicional, cabe resumir estas diferenças em um mesmo lugar. A tabela 6.1 expõe, em linhas gerais, as diferenças entre o que se entende por metodologia tradicional e a metodologia proposta nesta pesquisa.

Tabela 6.1: Comparação entre o que se considera como metodologia tradicional e a metodologia do PBL apoiada na TCC abordada nesta pesquisa.

Ente Didático(*)	Metod. Tradicional	PBL com TCC
Papel do Professor/Tutor	Expor de maneira clara e acessível o conteúdo, manter a disciplina dos alunos, avaliar o rendimento.	Propor situações cuja pesquisa permita o domínio gradual dos conceitos que se quer trabalhar. Orientar tal pesquisa, prover materiais de apoio e avaliar o nível de domínio alcançado do CC estudado.
Papel do aluno	Receber o conteúdo, compreende-lo pelo estudo das anotações de aula, do material de apoio e da resolução de exercícios. Verificar seu aprendizado por meio de questões de prova.	Assumir atitude de pesquisador para resolver os problemas provenientes da situação principal, propor e buscar soluções de novas questões norteadoras, assumir o protagonismo de seu processo de domínio do CC trabalhado.
Aula expositiva	O Professor expõe o conteúdo levando em conta a lógica do próprio conteúdo.	O conteúdo raramente é exposto, somente quando o Professor/Tutor nota uma barreira intransponível para os alunos.
Trabalho em grupo	Visa a resolução conjunta de perguntas propostas pelo Professor ou a pesquisa de um determinado tema.	Visa a resolução colaborativa de problemas. Usado para a construção do conteúdo por meio da busca de respostas para uma situação inicial e de questões norteadoras acerca desta situação.
Material de apoio	Descrevem o conteúdo, apresentam exemplos de exercícios e aplicações e propõem exercícios	Orientam a pesquisa, estruturam o conteúdo, ampliam conceitos
Lista de exercícios	Amplamente utilizadas para fixação e alguma ampliação do conteúdo, normalmente ao final do processo	Visam ampliar as situações que justificam os conceitos trabalhados, aparecem no decorrer do processo.

(*)Pessoa, aspecto ou recurso próprio do ensino.

A seção 7.1, do próximo capítulo, discorre sobre as experiências vividas durante a aplicação da metodologia descrita acima, procurando descrever o ambiente de ensino e seus atores, na medida em que os objetivos desta pesquisa exigem.

As demais seções do capítulo seguinte descrevem os resultados obtidos acerca do rendimento e do nível de domínio relativos ao CC dos THMP durante a aplicação do PBL apoiado pela TCC proposto nesta pesquisa.

Capítulo 7

Resultados

Os instrumentos descritos na seção 4.2 geraram dados cuja análise é apresentada a seguir.

7.1 Diário de Classe

Nesta seção serão analisados os resultados de cada um dos quatro encontros da aplicação definitiva (cuja metodologia está descrita no capítulo 6), tendo por base os dados do diário de classe criado pelo professor/tutor, logo após cada encontro, em arquivos de áudio, ressaltando as características marcantes da atividade. Os diários de classe referem-se às três turmas, constituintes da amostra, que participaram das atividades do PBL: PSC-A, PSC-C e NUT-A.

Este levantamento, ancorado nos registros do diário de classe, são de cunho interpretativo e descritivo e seu desfecho produz uma ideia particular e pontual da realidade experimentada pelo professor/tutor e têm o caráter de resultado de pesquisa, discutidos no capítulo 8.

Em termos metodológicos, os resultados aqui apresentados foram extraídos da seguinte maneira: os áudios contendo os diários de classe (mais de um áudio por encontro,

que juntos formam um diário de classe) foram ouvidos e a sua mensagem essencial anotada. De posse destas ideias essenciais, redigiu-se a primeira versão do texto com os resultados daquele encontro. Assim foi feito para todos os quatro encontros e o texto compilatório dos textos dos encontros é apresentado mais adiante.

O diário de classe, tal como apregoado na subseção 4.2.1, cumpriu um papel crucial no entendimento do processo de desenvolvimento do PBL, bem como na identificação de potenciais esquemas de pensamento utilizados pelos estudantes, mesmo que o diário de classe esteja mais propenso à descrição e entendimento da aula como um todo, desenvolvida em conjunto com todos os alunos, a classe.

O diário de classe reflete as impressões do professor/tutor, tão importantes ao fazer pedagógico de sala de aula e para esta pesquisa em Didática. Neste sentido, as reflexões aqui expostas procuram trazer uma análise geral do processo desenvolvido, tentando identificar fenômenos ocorridos nas três turmas que vivenciaram o PBL, e que possam subsidiar um fazer didático mais abrangente e representativo, no futuro, bem como aos objetivos desta pesquisa, em particular.

Assim, a análise que se segue atém-se às sutilezas captadas pela sensibilidade do professor/tutor, que registrou suas impressões no diário de classe e as analisou em momento futuro.

Dado que este instrumento tem forte caráter subjetivo, estabeleceu-se duas categorias de análise, norteadoras desta complexa coleta de dados: 'ensino' e 'aluno'.

A categoria 'ensino' dá conta das impressões do professor/tutor em relação ao processo de ensino, da metodologia utilizada e suas nuances e possíveis modificações, é a visão do professor/tutor quanto ao desenvolvimento da metodologia de ensino utilizada.

Por sua vez, a categoria 'aluno' trata das impressões do professor/tutor quanto à repercussão da metodologia nos alunos, ou seja, por meio dos sinais explícitos pelos alunos, coletou-se informações subjetivas acerca do sentimentos e do aprendizado dos

alunos. Esta categoria levanta dados por meio das manifestações verbais e gestuais e as coloca na perspectiva de turma (classe de alunos). É conveniente salientar que análises mais individualizadas, com respostas dos alunos dadas por escrito, estão disponíveis na subseção 7.3.2.

No texto que segue, as duas categorias são apresentadas de forma mesclada, pois é desta forma que o fazer pedagógico de sala de aula se dá, mas a escrita procurou deixar clara as categorias, tanto quanto possível.

7.1.1 Descrição do Diário de Classe

No primeiro encontro, os alunos das três turmas mostraram familiaridade com a base de dados utilizada (mesmas da aplicação prévia do PBL, descritas na seção 5.2: ESV para as turmas PSC-A e PSC-C e Bariátricos para a turma NUT-A), mas tiveram certa dificuldade em entenderem a diferença para as novas situações propostas pois julgavam que poderiam respondê-las com a análise estatística que entregaram durante o conteúdo de Dispersão (descrito na seção 5.2). Neste sentido, orientou-se que dessem uma atenção maior às questões norteadoras prévias produzidas pelo professor/tutor (anexos F e G), a fim de perceberem o foco da nova teoria que estavam iniciando a aprender. Este tipo de orientação é típica do PBL pois preserva o caráter investigativo no aluno, orientando o processo sem dar-lhe alguma resposta pronta.

Embora as turmas já tivessem experimentado o PBL, a formulação de suas próprias questões norteadoras aconteceu de maneira informal, mais como dúvidas pessoais do que como questões a pesquisar. Talvez por este motivo, a atividade foi muito confundida com um trabalho em grupo típico onde as questões são respondidas e entregues pelo grupo, sem a preocupação de desenvolver um conteúdo novo. Neste ponto ressalta-se a importância do professor/tutor em instigar a produção de novas questões norteadoras pois, de forma espontânea, nem sempre surgem.

Neste sentido, o trabalho em grupo típico do PBL, onde há uma discussão acerca

do tema, a geração de questões norteadoras, a divisão de tarefas e a troca de resultados particulares de pesquisa, aconteceu de maneira intuitiva, pouco estruturada e com questões de cunho mecânico, do tipo “como se calcula” mais do que do tipo “o que significa”. Contudo, o trabalho em grupo desenvolvido manteve o caráter investigativo exigido pelo PBL, muito pela ação do professor/tutor ao instigar esta atitude nos grupos.

Durante a evolução do processo, percebeu-se um distanciamento em relação à metodologia tradicional no que se refere ao trabalho em grupo, no sentido de que a investigação, visando a resolução de problemas, ficou evidente pela necessidade de avanço para além das questões norteadoras e do material de apoio entregues pelo professor/tutor, obrigando que o grupo fizesse um movimento para além de responder a determinadas questões ou pesquisar em torno de um tema, típicos do trabalho em grupo na metodologia tradicional.

Em conversa ao final do primeiro encontro, foram tratadas questões referentes às primeiras quatro perguntas disponibilizadas pelo professor/tutor. A questão número um foi facilmente respondida com o auxílio do Excel, mostrando que a ferramenta para o cálculo da média estava dominada.

Contudo, as próximas questões referiam-se à noção intuitiva de probabilidade e foram respondidas mais com “sim” ou “não” do que com ponderações maiores. Neste ponto, o professor/tutor fomentou a discussão acerca destas questões para fazer surgir a necessidade de cálculos que melhorassem a noção intuitiva que haviam mostrado. Por exemplo, na questão número três, supõe-se que uma população tenha uma determinada média e pergunta-se se há alguma chance de se retirar uma amostra cuja média não seja a populacional. Neste exemplo, os alunos tenderam a responder somente que “sim, é possível”, sem a preocupação em estender o raciocínio à distribuição amostral da média, por exemplo.

Durante uma explicação do professor/tutor acerca da plausibilidade de obtenção de uma amostra oriunda de população com média 25, invocando o conhecimento previa-

mente estudado da distribuição amostral da média, um aluno ponderou o seguinte: Se é conhecida a média da população, qualquer amostra retirada dela é plausível.

Diante disso, é importante notar a relevância dos conhecimentos adquiridos antes dos THMP. A distribuição amostral da média (DAM) permite o cálculo da probabilidade de obtenção de uma determinada média amostral oriunda de uma população, neste caso, de média conhecida. Por este meio, pode-se determinar o quão plausível é a ocorrência de uma determinada média amostral a partir de uma população com média conhecida.

Mesmo havendo uma discussão sobre o pensamento probabilístico neste primeiro encontro, a falta de uma resposta precisa pareceu incomodar os alunos. Para eles, parece ser muito mais simples o cálculo da probabilidade, valor único que remete à ideia de precisão, do que ponderar sobre a incerteza subjacente ao número. Parece ser frustrante, para os alunos, desenvolver o processo probabilístico, com valores bem definidos e ter que ponderar sobre a incerteza do processo.

Com a discussão das primeiras quatro questões, parece ter havido um aumento na preocupação dos alunos em relação à atividade, no sentido de terem que dar conta de um conteúdo sem a explicação do professor. O desconforto de serem protagonistas na busca do conhecimento resultou em uma preocupação com a resolução da questão número cinco, estruturada para os THMP. Desta forma, o planejamento dos grupos para o trabalho fora da sala de aula, crucial para o bom andamento da atividade, pareceu satisfatório na medida em que os grupos, agora, pesquisaram para além da estrutura dada, preocupando-se em pormenorizá-la e preencher suas lacunas por meio do trabalho investigativo, em grupo.

Também transpareceu a preocupação dos alunos em relação às notas e o professor/tutor teve a impressão que a motivação gerada pela situação já não era preponderante em relação à motivação que a necessidade de boa nota produzia e os alunos estavam atribuindo a nota às respostas das questões, o que não era verdade pois as questões somente faziam parte da estruturação do PBL e não da avaliação.

No segundo encontro, em geral os grupos mostraram um bom nível de trabalho extraclasse pois houve muitas perguntas acerca das questões dentro da questão número cinco. Os grupos se deram conta de que tais questões não tinham resposta simples e direta, dependentes de um único conceito.

Houve perguntas sobre a obrigatoriedade de se trabalhar com os materiais de apoio e em que ordem os materiais deveriam ser trabalhados. O professor/tutor orientou que os materiais de apoio não eram obrigatórios, senão um auxílio à pesquisa e que não havia ordem para consumi-los.

Em geral, os alunos pareceram motivados porém angustiados e o segundo encontro exigiu muito do professor/tutor, tanto no esclarecimento de dúvidas em relação aos conceitos quanto na orientação de um bom caminho a seguir para otimizar a pesquisa dos grupos.

Neste segundo encontro surgiu a preocupação com os passos dos THMP, o que era esperado dado o tipo de perguntas da questão cinco. O professor/tutor chamou atenção para os quatro passos dos THMP, destacando a abordagem do p-valor: definição das hipóteses, estabelecimento do nível de confiança, cálculo do p-valor e tomada de decisão. Este enfoque procurou limitar a gama de conteúdos a serem pesquisados neste ponto, visando acelerar o domínio procedimental dos THMP.

Ficou clara a necessidade de maior estruturação do processo neste ponto, então, o professor/tutor proferiu uma explicação acerca da ideia geral dos testes de hipóteses utilizando o exemplo da verificação da autenticidade de uma moeda usada em um jogo de “cara ou coroa”. Neste exemplo, escolhe-se uma moeda para analisar, pondera-se sobre uma hipótese nula possível, por exemplo: H_0 : a moeda é honesta. Logo, uma hipótese alternativa possível é H_1 : a moeda não é honesta. Em seguida, cria-se um teste onde se lança a moeda, digamos vinte vezes, anotando a quantidade de caras obtidas. É esperado que saiam 10 caras, mas, por conta da aleatoriedade, pode sair um número diferente de dez sem que a moeda seja falsa, por isto, define-se o que seria um nível

de diferença de dez aceitável, antes do qual, não haja desconfiança da hipótese nula de honestidade da moeda.

Pois bem, na turma PSC-C, escolheu-se uma quantidade de caras entre cinco e quinze como zona de aceitação da hipótese nula e uma amostra de vinte lançamentos gerou 13 caras. Decidiu-se, então, que a amostragem não ajudou a refutar a hipótese nula em benefício da alternativa.

Este exemplo precisou ser trabalhado mais de uma vez, reforçando seus detalhes e fazendo perguntas aos Alunos, do tipo: foi provado que a moeda é verdadeira? Onde parece ter ficado claro para a maioria de que o teste não é uma demonstração.

O exemplo mostrou-se útil pois a ideia de erro do tipo um e dois, nível de confiança e tomada de decisão também foram trabalhados, deixando ainda em aberta as possibilidades de pesquisa para os Alunos.

É válido ressaltar que este exemplo não estava previsto na metodologia de ensino descrita no capítulo 6, mas foi uma necessidade observada pelo professor/tutor a partir da prática pedagógica, fazendo perceber que era preciso um nível de estruturação maior do que o planejado.

Uma questão comum, explicitada pelo exemplo é a dificuldade do estabelecimento da hipótese nula, no sentido dela ser uma afirmação e, por isto, ser interpretada como certeza. Esta dificuldade é aumentada no estabelecimento da hipótese alternativa, antagônica da nula. É motivo de confusão afirmar algo em um momento e o seu oposto em outro (a moeda é honesta ou não). É preciso ter a sensibilidade de identificação deste problema e colocar as hipóteses estatísticas num caráter de suposição, premissa.

Uma OD surgiu com a pergunta: o nível de significância é uma porcentagem e o Sr. não usou porcentagem pra estabelecer o intervalo entre 5 e 15. O professor/tutor respondeu que o nível de significância tem a ver com uma probabilidade de ocorrência de caras nas vinte jogadas. Pra isso, é preciso conhecer todas as probabilidades de ocorrência de caras, de nenhuma até vinte, nos vinte lançamentos. O nome disto é distribuição de

probabilidades para a proporção e vocês usarão a distribuição de probabilidades para a média, que seguem dois padrões, em nosso caso, o dado pela distribuição Z e outro pela distribuição T . O invariante operatório, do tipo conceito em ação, subjacente ao questionamento do aluno pode ser: “o nível de significância é uma porcentagem”. A abordagem do professor/tutor foi no sentido de expandir a ZDP do aluno para alcançar as distribuições de frequências que irão ser úteis no desenvolvimento dos THMP.

Outra OD suscitada pelo exemplo iniciou com a afirmação: Se o teste não prova que a moeda é falsa, não adianta nada rejeitar a hipótese nula. O professor/tutor argumentou que o teste reforça ou não uma suposição inicial, por meio da análise de uma amostra e que isto é útil, mas não é uma prova. O invariante operatório que pode estar em uso neste caso parece ser do tipo conceito em ação: “A incapacidade de demonstração torna os testes de hipóteses inúteis” e a ruptura proposta vai no sentido de que utilidade não significa, necessariamente, demonstração e que a estatística aumenta a informação acerca do problema que se tem, na tentativa de aumentar o conhecimento sobre o assunto em questão.

O terceiro encontro foi o momento de compartilhamento das respostas dos grupos acerca da situação principal. Em resumo, os grupos apresentaram a resolução da questão usando o processo dos THMP e a abordagem do p-valor, com baixa interpretação dos resultados. Os grupos não acrescentaram um raciocínio mais amplo do que simplesmente rejeitarem H_0 . Pareceu claro que a parte mecânica estava, se não totalmente, grandemente dominada e que a interpretação subjacente não era encarada como de suma importância pelos Alunos. Seguiu-se, então, uma exposição do professor/tutor acerca da interpretação dos resultados.

No terceiro encontro, o processo dos THMP estava melhor dominado, com a parte mecânica bastante avançada, principalmente com o cálculo do p-valor, mas ainda carente de ajustes no que se referia ao estabelecimento da zona de rejeição de H_0 e o uso das tabelas Z e T .

O professor/tutor sentiu a necessidade de produzir um vídeo explicativo do uso das tabelas Z e T, o que foi feito e disponibilizado aos Alunos entre o terceiro e quarto encontros. Os vídeos tiveram uma aceitação não mais do que razoável, muito por conta da facilidade do uso do p-valor em comparação como o uso das tabelas Z e T.

Ao final do terceiro encontro, algumas dúvidas acerca das questões extras (anexo H) foram sanadas. Uma OD foi identificada por meio da pergunta: Se eu rejeitei a hipótese nula, é porque ela é falsa. O professor/tutor respondeu com o questionamento: Rejeitamos a hipótese nula baseados em uma amostra. Se tirássemos outra amostra, rejeitaríamos também? E o aluno ponderou: Claro, o teste provou que vale pra todas. Daí seguiu-se uma discussão acerca da distribuição amostral da média, no intuito de contrapor o invariante operatório, do tipo teorema em ação, que aparentemente o aluno lançou mão: “Se o teste rejeita a hipótese nula, então a hipótese é falsa”.

O quarto encontro foi dedicado ao esclarecimento de dúvidas individuais acerca das questões extras, onde notou-se uma heterogeneidade entre os alunos. Alguns pareciam estar engajados e conhecedores do processo e outros estavam tentando “decorar” os passos dos THMP na última hora, antes da avaliação.

Ao final do quarto encontro foram aplicadas as duas questões abertas como descrito no capítulo 6. A análise das respostas dos alunos a estas questões está na seção 7.3.

Durante a avaliação do quinto encontro, uma Aluna estava em dúvida entre duas alternativas, uma dizia que nos THMP pressupõe-se que a hipótese nula é verdadeira e na outra dizia que os THMP comprovam que a hipótese nula é verdadeira. O professor/tutor respondeu que, por definição, o p-valor é uma probabilidade, e fez a seguinte pergunta: quando se calcula uma probabilidade se está demonstrando que aquilo é correto, ou não? Onde a Aluna respondeu que não é uma demonstração, mas continuou sem saber qual das alternativas escolher. Este fato suscita a ideia de que os Alunos vêem os THMP como uma demonstração, afinal, se está rejeitando ou não a hipótese nula, o que Eles vêem como a demonstração ou não de sua veracidade. Isto tanto é verdade

que, mesmo a Aluna raciocinando corretamente, não teve a convicção necessária para desconstruir esta impressão de determinismo que o teste traz consigo.

Um invariante operatório é extremamente difícil de ser identificado e mesmo seus indícios, profícuos para as OD, nem sempre são passíveis de identificação. No caso acima, o professor/tutor não conseguiu identificar um possível invariante operatório, por isso, a situação não configurou uma OD.

O professor/tutor perguntou a um aluno se ele estava gostando da metodologia. Ele respondeu que aprendia melhor com a explicação do professor. Este mesmo aluno já havia demonstrado um bom nível analítico durante as aulas, inclusive trazendo interpretações interessantes acerca da impossibilidade de se provar uma afirmação populacional com a retirada de uma única amostra. Mesmo o aluno tendo alcançado conhecimentos que não foram “dados” pelo professor/tutor, Ele desconfia do seu progresso.

A angústia gerada pelo estudo mais individualizado, autônomo, fica evidente, mesmo quando os alunos estão em grupos. Existe um sentimento de incompletude dos conceitos que estão sendo estudados autonomamente. É curioso pensar que este sentimento salutar é minimizado no processo tradicional de ensino onde o professor traz “todo o conteúdo”, dando ares de completude ao processo.

Cabe lembrar que o capítulo seis e a seção 7.1 são complementares no sentido da descrição da metodologia de ensino empregada nesta pesquisa. Visto que o PBL parte de um problema clínico real (situação principal, nesta pesquisa) e evolui conforme as pesquisas dos alunos se desenvolvem, com a orientação do professor/tutor, as novas situações geradas pelos próprios alunos retroalimentam o processo do PBL, por isto não puderam ser previstas no capítulo seis, mas foram descritas acima.

Neste sentido, esta pesquisa trouxe estruturações no processo do PBL, motivadas pelo ambiente de ensino e pela TCC. Tais estruturações, na forma de disponibilização de materiais de apoio e no conteúdo específico dos THMP, não inviabilizam o PBL, pelo contrário, visaram torna-lo factível num contexto de ensino tradicional.

7.1.2 Resultados do Diário de Classe

O Diário de Classe permite o levantamento de importantes resultados de pesquisa, entretanto, cabe novamente ressaltar que, a esta altura, tais dados referem-se às impressões do professor/tutor, na tentativa de produzir um quadro de conhecimento mais completo acerca do processo de ensino empreendido. Tal quadro ganhará contornos mais determinísticos quando forem juntados os resultados dos outros instrumentos utilizados nesta pesquisa. Por hora, ficam as tabelas [7.1](#) e [7.2](#), como resumo de tais impressões.

Tabela 7.1: Resumo dos resultados auferidos pelo Diário de Classe.

Ente Didático(*)	'Ensino'	'Aluno'
Atividades Prévias	Aceleraram a aplicação das atividades definitivas.	Aluno familiarizado com a base de dados.
Questões norteadoras	Tendem a não nascerem por ampla discussão coletiva. A estruturação dada pelo professor/tutor auxilia neste processo	Dúvidas individuais comumente tornam-se questões norteadoras.
Trabalho em grupo	No início, a estrutura tradicional de trabalho em grupo pareceu manter-se para as questões norteadoras prévias. Na sequência, a necessidade de resolverem problemas abertos e a orientação do professor/tutor, caracterizou o trabalho em grupo como típico do PBL.	O aluno teve a tendência de responder às questões norteadoras sem a preocupação de ensinar os colegas do grupo. A necessidade do envolvimento mais propositivo do que reativo incomodou alguns alunos, persistindo este incômodo até o final das atividades.
Trabalho individual	Os grupos demonstraram ansiedade ao se depararem com problemas abertos e, em geral, não associaram rapidamente, que o material disponibilizado não tinha um fim em si mesmo.	Houve uma insegurança a respeito do papel a cumprir no grupo, normalmente minorizada pela atribuição de tarefa a realizar.
Recurso computacional	As atividades prévias prepararam minimamente os alunos para o uso dos computadores, porém, individualmente, ainda se observava alunos dependentes dos demais no manuseio do Excel.	A ansiedade estatística parece ser compartilhada com uma ansiedade computacional. Muitos alunos tinham medo do computador. A destreza que se observa no manuseio de navegadores e redes sociais não transmitiu a mesma naturalidade para o uso de softwares como Word e Excel, para muitos alunos.
Aula expositiva	Houve dúvidas gerais que impuseram o uso de aula expositiva para além da explicação das atividades, envolvendo outros exemplos e expondo fragilidades notadas nos alunos pelo professor/tutor. Não fosse isto, a metodologia cobriria um tempo demasiadamente grande.	Pareceu ser reconfortante a alguns alunos que não estavam acostumados com o protagonismo em seu aprendizado. Por ter partido da necessidade dos alunos, as explicações sanaram dúvidas particulares, melhorando o ânimo de alguns alunos.

(*)Pessoa, aspecto ou recurso próprio do ensino.

Tabela 7.2: Resumo dos resultados auferidos pelo Diário de Classe (continuação).

Ente Didático(*)	'Ensino'	'Aluno'
Pensamento Probabilístico	As questões acerca da probabilidade precisaram ser incitadas pois a variabilidade e a aleatoriedade não foram abordadas naturalmente	Os alunos não dissociaram o valor numérico da probabilidade com a ocorrência de um evento particular.
Exemplo da moeda	Apresentou uma perspectiva estruturada e de mais fácil adequação à situação da área da saúde.	Sanou algumas dúvidas dos alunos e pareceu lhes diminuir a ansiedade.
Estruturação dos THMP	Estruturar o conteúdo dos THMP em quatro partes foi importante para os alunos pois lhes facilitou o processo de pesquisa deste conteúdo.	Ficou mais fácil a separação de demandas e a pesquisa individual do aluno, bem como a apresentação de seus achados ao grupo.
Motivação	Foi difícil separar sinais de motivação dos sinais de empenho resiliente na realização da tarefa. A mesma dificuldade apareceu para diferenciar se o trabalho era motivado pela situação problema ou pela nota a ser alcançada.	Alunos que tiveram um posicionamento mais participativo dentro do grupo aparentaram maior motivação. Alunos que tomaram um papel de liderança destacaram-se quanto à motivação.
Oportunidade Didática	O grupo foi gerador de Oportunidades Didáticas potencialmente aplicáveis a grande parte dos alunos. Mesmo reconhecendo que os invariantes operatórios são diferentes entre os alunos, parece ser útil desenvolver linhas de raciocínio individualizada na presença do grupo ou de toda a turma	A liberdade que a metodologia proporciona ao descentralizar a busca do conhecimento do professor, libera tempo para que o professor/tutor crie um vínculo intelectual com um aluno em particular por meio da Oportunidade Didática

(*)Pessoa, aspecto ou recurso próprio do ensino.

Deve-se salientar que o instrumento Diário de Classe deu delineamentos gerais da experiência vivida que, embora não generalizáveis por causa do forte caráter particular desta aplicação, são ricos em informações relevantes ao fazer pedagógico.

Estando apresentado o contexto qualitativo vivenciado, passa-se à análise do rendimento e do nível de conquista do CC dos THMP, auferido pelos estudantes durante o processo.

7.2 Análise do Rendimento

O instrumento fechado, composto por questões de múltipla escolha (itens), está descrito na subseção 4.2.2 e os resultados gerados por ele são descritos a seguir. A análise do instrumento fechado também pode ser encontrada em Fioravanti et al. (2020).

O rendimento dos estudantes (score) tem as medidas resumo apresentadas na figura 7.1.

		Estadística
Escore	Média	6,56
	Mediana	6,00
	Desvio Padrão	1,960
	Mínimo	2
	Máximo	10

Figura 7.1: Medidas resumo para a variável score. (Gerado por SPSS)

De forma geral, o rendimento dos estudantes é considerado razoável se levarmos em conta a média, que é suficiente para aprovação na instituição (cuja média para aprovação é de seis pontos) onde a pesquisa está sendo desenvolvida e o desvio padrão que determina um coeficiente de variação de 30% que é alto, porém comum para esta instituição¹.

A distribuição dos escores dos 99 alunos, com a respectiva curva normal, é mostrada na Figura 7.2.

¹Embora não haja estudo que permita a comparação destes resultados, acredita-se, pela experiência prévia, que estes resultados sejam, de fato, típicos desta instituição.

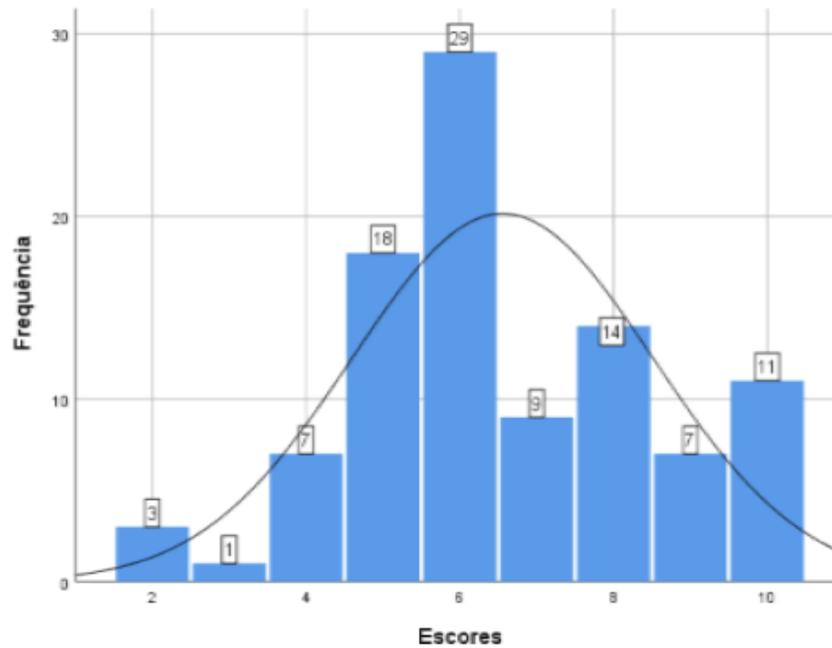


Figura 7.2: Distribuição de frequências de escore, média, desvio padrão, quantidade de elementos (N) e distribuição normal correspondente. (Gerado no SPSS)

Percebe-se, pela figura 7.2, que os itens 3 e 7 possuem frequências bem menores do que as esperadas para a distribuição normal subjacente, enquanto que os itens 6 e 10 possuem frequências bem maiores. Neste caso, a imagem parece mostrar que não há adequação da distribuição dos escores com a curva normal correspondente.

A Figura 7.3 apresenta o boxplot dos escores, com os quartis, a média (6,56) e o desvio padrão (1,95), o escore mínimo (2) e o máximo (10).

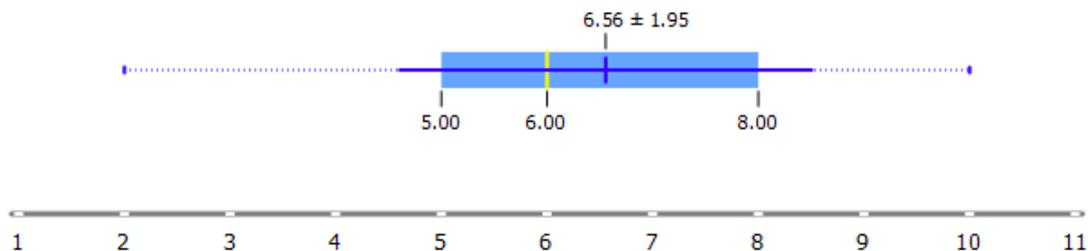


Figura 7.3: *Boxplot* para a variável escore com média, desvio padrão, quartis e *whiskers*. (Gerado no Orange Canvas))

Tabela 7.3: Medidas da TCT por item. (Gerado por Excel e R, pacote ltm)

Item	Índice de dificuldade	Índice de discriminação	Coef. Ponto Bisserial	Alfa de Cronbach (geral: 0,5437) Excluindo os itens
Q1	86%	0,21	0,28	0,54
Q2	40%	0,35	0,41	0,54
Q3	62%	0,62	0,53	0,49
Q4	45%	0,54	0,52	0,5
Q5	51%	0,5	0,48	0,51
Q6	71%	0,46	0,41	0,53
Q7	90%	0,28	0,42	0,51
Q8	73%	0,63	0,47	0,51
Q9	49%	0,7	0,57	0,48
Q10	89%	0,24	0,28	0,54

O teste de aderência de qui-quadrado para normalidade apresentou um valor calculado igual a 16,03, contra um valor tabelado de 12,59, a um nível de 5% de significância, levando à rejeição da hipótese de normalidade dos dados para a variável score. O teste também indicou que os valores observados mais discrepantes dos valores esperados, que colaboraram para a rejeição da normalidade, ocorreram nos escores 6, 7 e 10, corroborando com a inspeção visual da Figura 1 que indicava estes escores com forte discrepância com a curva normal subjacente, a menos do escore 3, que o teste qui-quadrado não identificou como discrepante ao valor esperado para a curva normal. Também o teste de normalidade de Shapiro-Wilk, ao mesmo nível de significância de 5%, não permitiu pressupor que os escores provenham de população normalmente distribuída ($p < 0,001$).

Por todos estes motivos contra a pressuposição de normalidade dos dados de score, a correlação bisserial foi substituída pela correlação ponto-bisserial, gerando a Tabela 7.3, juntamente com outras medidas da TCT.

O Alfa de Cronbach geral (0,5437) não é alto, mesmo que o valor 0,6 seja considerável aceitável para alguns cenários de pesquisa, levando-se em conta que instrumentos grandes tendem a ter valores de α maiores (Maroco e Garcia-Marques, 2006), o que

não é o caso do instrumento aqui analisado. A exclusão de itens também não gerou uma elevação na confiabilidade, sendo desnecessária a redução do número de itens do instrumento.

O item Q2 teve a menor proporção de acertos (40%), mesmo assim, considera-se o item acessível aos estudantes, visto que obteve um percentual considerável de acertos. Ao mesmo tempo, os itens Q1, Q7 e Q10 alcançaram mais de 86% de acertos, sendo considerados os mais fáceis.

A Figura 7.4 mostra o gráfico com o escore total em relação com a proporção de acertos dos itens Q1 (1), Q7 (7) e Q10 (d), revelando que mesmo os estudantes com baixos escores já alcançam altas proporções de acertos para estes itens, indicando que eles não são capazes de distinguir alunos com alto rendimento dos com baixo rendimento. Observa-se que esses três itens possuem baixa discriminação (Tabela 7.3), o que geralmente acontece com itens muito fáceis ou muito difíceis.

No gráfico da Figura 7.4 também percebe-se um padrão diferente no comportamento das curvas para o escore igual a três. Isso se deve ao fato que somente um aluno obteve escore 3 (Figura 7.2), o qual acertou os itens Q1 e Q7 e errou o item Q10, causando essa distorção no comportamento da curva. Algo semelhante ocorre com o escore 2, obtido por apenas 3 alunos: todos acertaram o item Q10 e erraram o item Q7 e apenas um acertou o item Q1 (33,3% de respostas corretas).

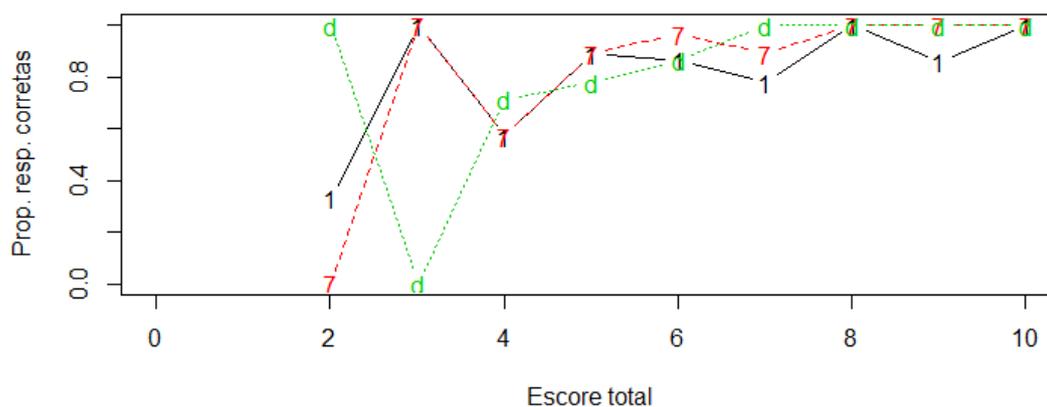


Figura 7.4: Gráfico do escore versus a proporção de acertos por item. O número 1 refere-se ao item Q1, o 7 ao Q7 e a letra d indica o item Q10. (Gerado pelo pacote 'ltm' do R)

Por outro lado, os itens Q2, Q4 e Q9 (menores índices de dificuldade, por isto, mais difíceis) só tiveram níveis altos de acertos para alunos com um rendimento, em geral, maior ou igual a sete (Figura 7.5), fazendo parecer que os itens são capazes de separar os candidatos com alto rendimento dos demais. Resultado reforçado pelos índices de discriminação altos dos itens Q4 e Q9 (0,54 e 0,70, respectivamente), ao contrário do item Q2 (0,35). Da mesma forma, o coeficiente ponto bisserial de Q4 e Q9 foram altos (0,52 e 0,57), ao contrário do item Q2 (0,41). Assim, parece que Q2, embora seja o mais difícil, acertá-lo não indica necessariamente bom rendimento no instrumento, enfraquecendo a capacidade do item de separar baixos rendimentos de altos rendimentos, ao contrário dos itens Q4 e Q9, também considerados difíceis. Quanto ao item Q9, poderia ter obtido um valor maior no índice de discriminação se não fosse pelo acerto, possivelmente ao acaso, do único aluno com escore 3 (Figura 7.2), causando uma mudança no comportamento da curva do item Q9 no escore 3.

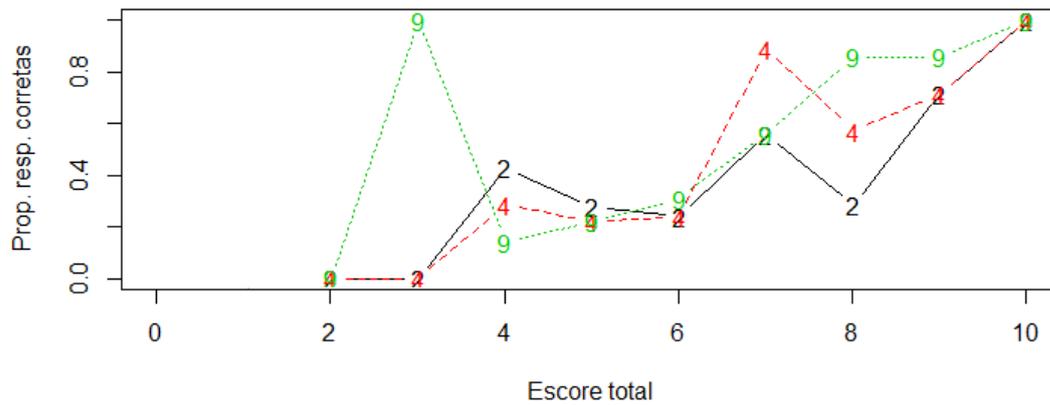


Figura 7.5: Gráfico do escore versus a proporção de acertos por item. O número 2 refere-se ao item Q2, o 4 ao Q4 e o 9 ao Q9. (Gerado pelo pacote 'ltm' do R)

Análise dos escores com estratificações

Passa-se, agora, a comparar os resultados dos escores por gênero, curso e participação no PBL, primeiramente, apresentando um resumo das médias considerando as variáveis estratificadoras, na Figura 7.6.

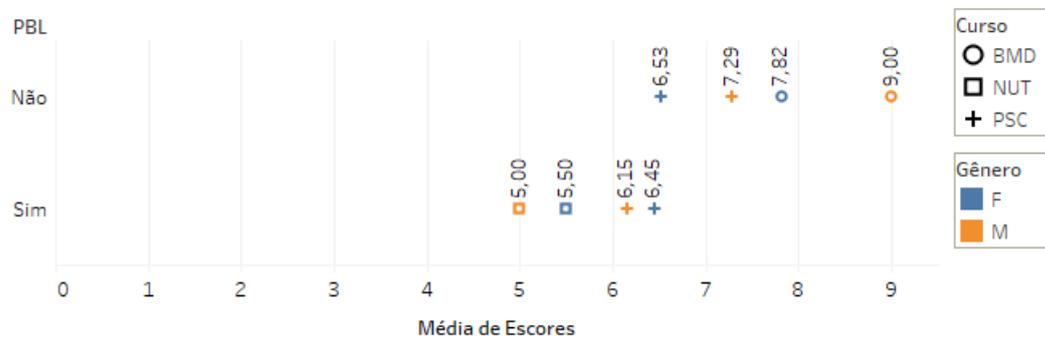


Figura 7.6: Médias dos escores separadas por participação no PBL, curso e gênero. (Gerado por Tableau)

As médias mais altas foram do curso de Biomedicina (nenhum dos 12 alunos participou do PBL), de ambos os gêneros, enquanto que as menores médias foram do curso de Nutrição (todos os 10 alunos participaram do PBL). O curso de Psicologia obteve médias intermediárias (51 alunos participaram e 26 não participaram do PBL). Assim, as médias mais altas foram dos que não participaram do PBL.

Estratificação dos escores por PBL separados por gênero

Dentre os 99 alunos da amostra, 76 (76,8%) declararam-se do gênero feminino e obtiveram uma média dos escores de 6,57 (2,12) , enquanto os 23 (23,2%) restantes, do gênero masculino, alcançaram um escore médio de 6,52 (1,31). As distribuições dos escores por gênero não são normalmente distribuídas segundo o teste de Shapiro-Wilk (p-valor para Feminino: 0,001 e p-valor para Masculino: 0,011) e o teste de Mann-Whitney não trouxe diferença significativa ($p = 0,939$) entre as médias de escores entre Feminino e Masculino.

Tomando-se os 76 indivíduos do gênero Feminino, 46 (60,5%) participaram do PBL e os 30 (39,5%) restantes não participaram do PBL.

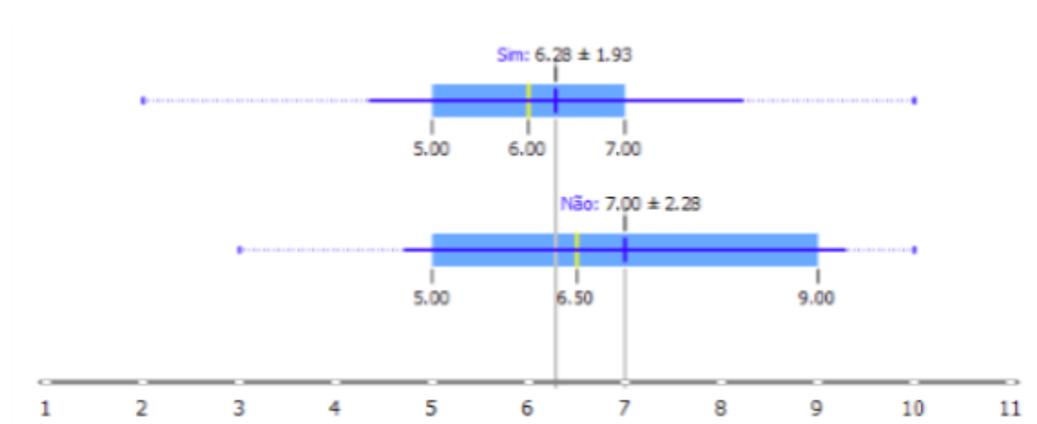


Figura 7.7: *Boxplots* dos escores com média, desvio padrão, quartis e whiskers, do gênero feminino que participaram (Sim) ou não participaram (Não) do PBL. (Gerado no Orange Canvas)

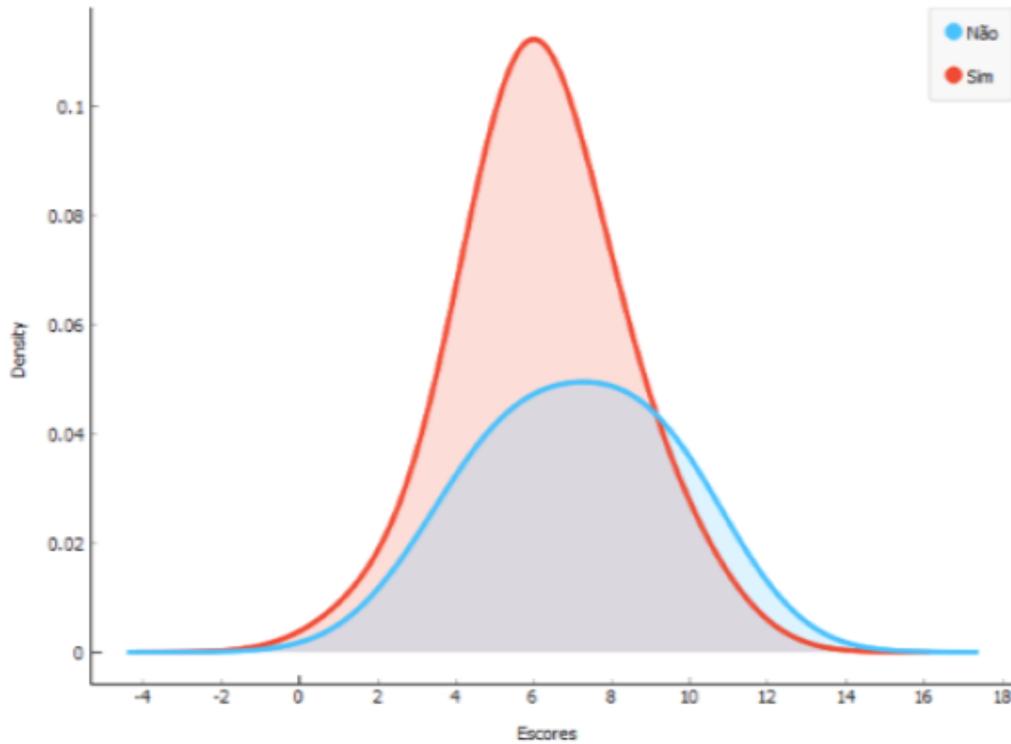


Figura 7.8: Curva de estimação da densidade de probabilidade núcleo gaussiano para os escores do gênero feminino, separados por participação no PBL. (Gerado por Orange Canvas)

Tanto a média quanto a variabilidade (desvio padrão) dos escores dos indivíduos do gênero Feminino que não participaram do PBL foram maiores do que os que participaram, entretanto, o teste de Shapiro-Wilk não permitiu pressupor que estes grupos provém de distribuições populacionais normais (p-valor: 0,007 e 0,016, participaram do PBL e não participaram, respectivamente), logo, usou-se o teste de Mann-Whitney para verificar-se que não houve diferença significativa entre as médias dos escores dos dois grupos (p-valor = 0,286).

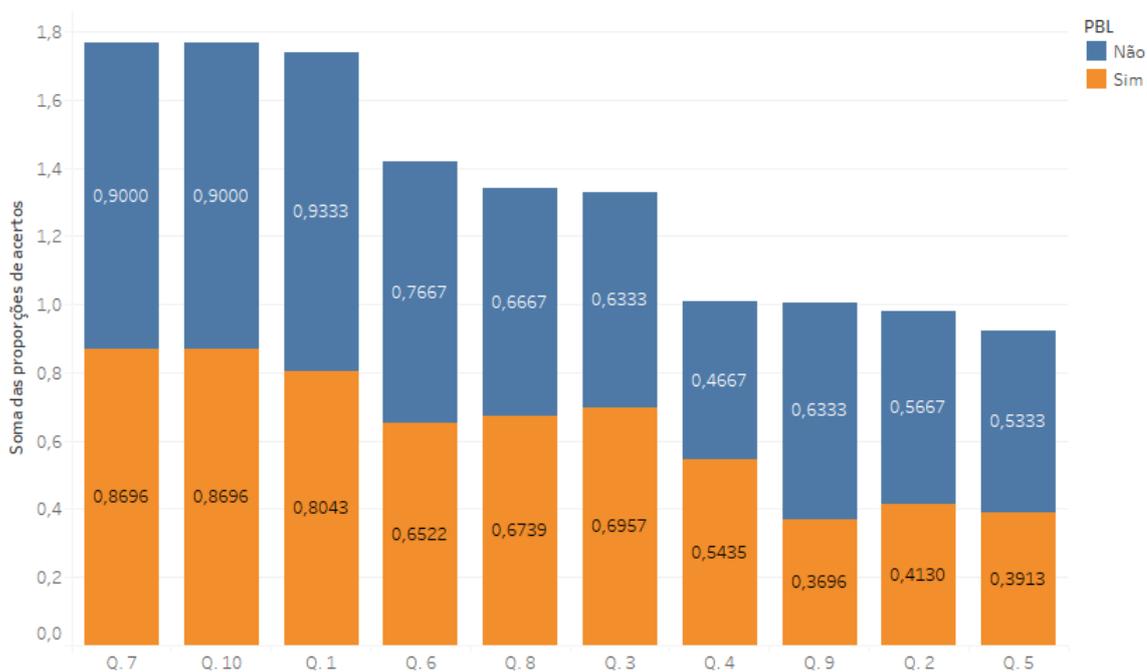


Figura 7.9: Proporções de acertos obtidos pelo gênero feminino, em cada questão (Gerado por Tableau).

Somente em três itens (Q8, Q3 e Q4) a média de acertos foi maior entre os indivíduos do gênero Feminino e o item Q9 teve um rendimento bastante baixo (0,3696) para quem participou do PBL, bem menor do que o rendimento das que não participaram (0,6333).

Já para os 23 indivíduos do gênero Masculino, 15 (65,2%) participaram do PBL e os 8 (34,8%) restantes não participaram do PBL.

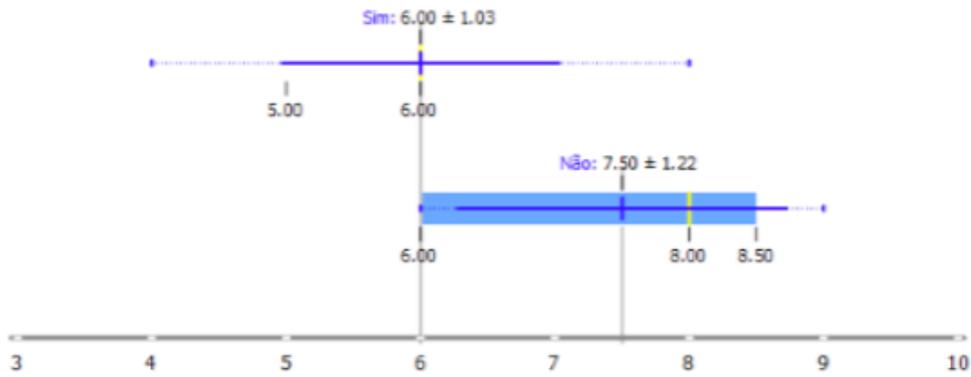


Figura 7.10: *Boxplots* dos escores com média, desvio padrão, quartis e whiskers, do gênero masculino que participaram (Sim) ou não participaram (Não) do PBL. (Gerado no Orange Canvas)

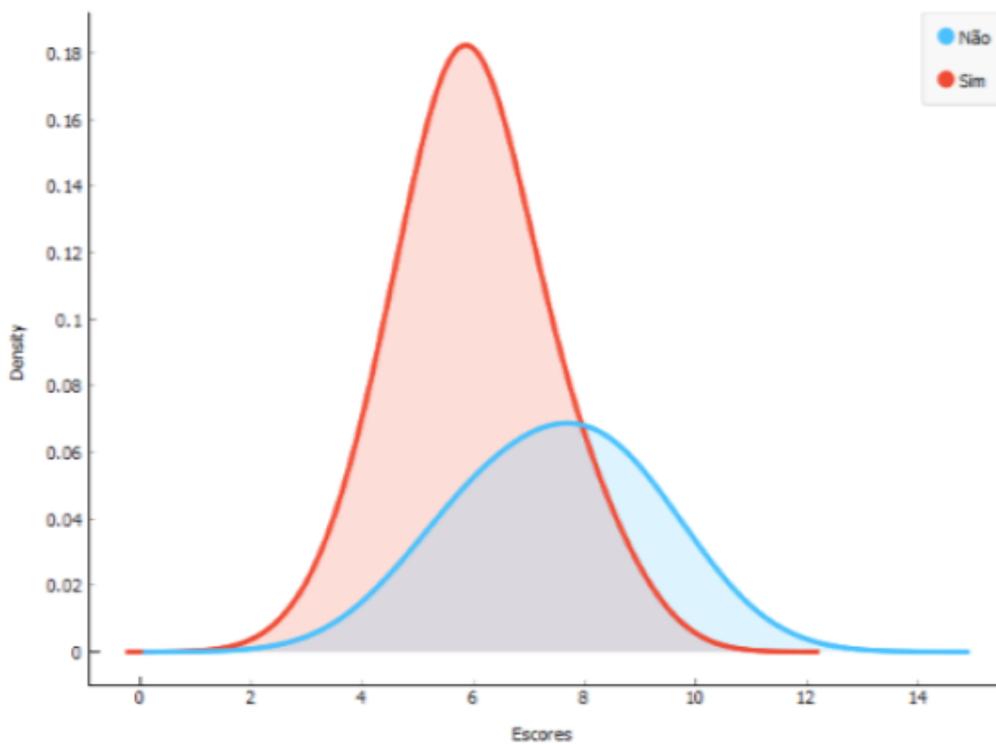


Figura 7.11: Curva de estimação da densidade de probabilidade núcleo gaussiano para os escores do gênero masculino, separados por participação no PBL. (Gerado por Orange Canvas)

Também para os Homens, a média e a variabilidade (desvio padrão) dos escores dos que não participaram do PBL foram maiores do que os que participaram, entretanto, o teste de Shapiro-Wilk não permitiu pressupor que estes grupos provém de distribuições populacionais normais (p-valor: 0,035 e 0,038, participaram e não participaram do PBL, respectivamente), levando ao uso do teste de Mann-Whitney para verificar-se que houve diferença significativa entre as médias dos escores dos dois grupos (p-valor = 0,019).

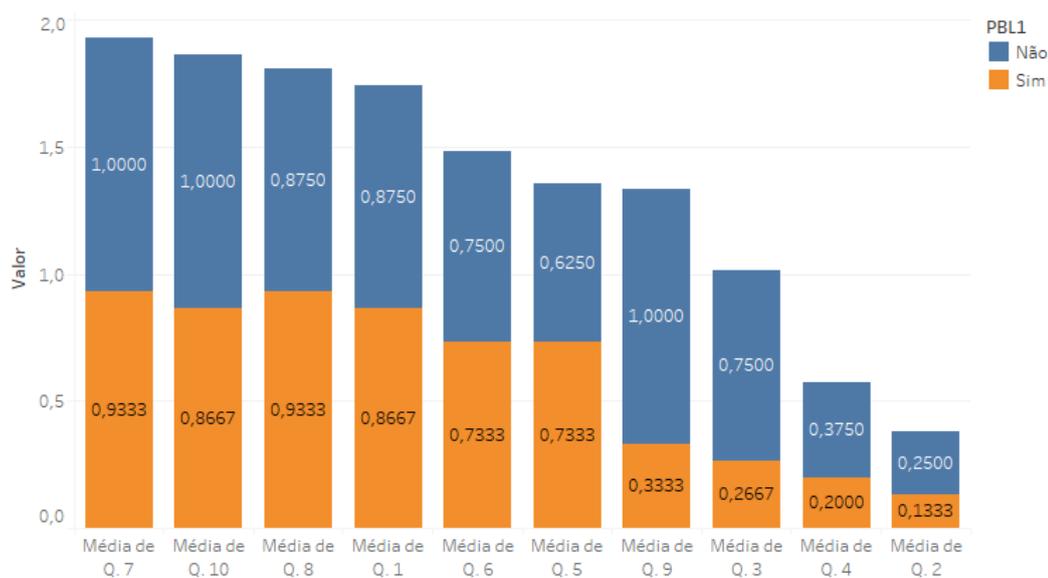


Figura 7.12: Proporções de acertos obtidos pelo gênero masculino, em cada questão (Gerado por Tableau).

Somente o item Q5 teve a média de acertos maior entre os indivíduos do gênero Masculino e é notório o baixo rendimento dos que participaram do PBL nos itens Q9, Q3, Q4 e Q2, sendo os itens Q4 e Q2 tendo baixos rendimentos em ambos os grupos.

Os itens Q9 e Q2 tiveram baixo rendimento tanto para Mulheres quanto para Homens que participaram do PBL sendo que os escores para Q9 dos que não participaram do PBL foi bem melhor do que os que participaram.

Estratificação dos escores por PBL separados por turma

Ao se estratificar os escores por turma, obtém-se as figuras abaixo que mostram

uma média maior para a turma da Biomedicina enquanto que a nutrição apresenta a menor média, entretanto a turma da Biomedicina também apresenta grande dispersão enquanto que a da nutrição, não. Lembrando que as turmas NUT-A, PSC-A e PSC-C participaram do PBL, ao contrário das demais.

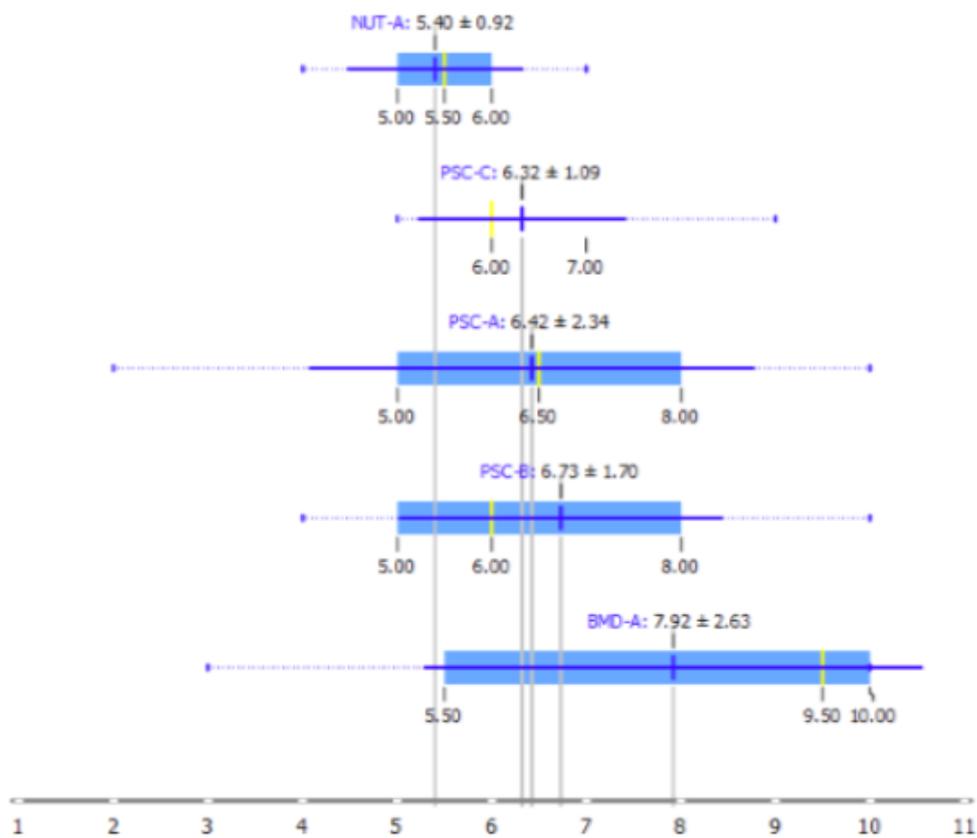


Figura 7.13: *Boxplots* dos escores com média, desvio padrão, quartis e whiskers, das diferentes turmas. (Gerado no Orange Canvas)

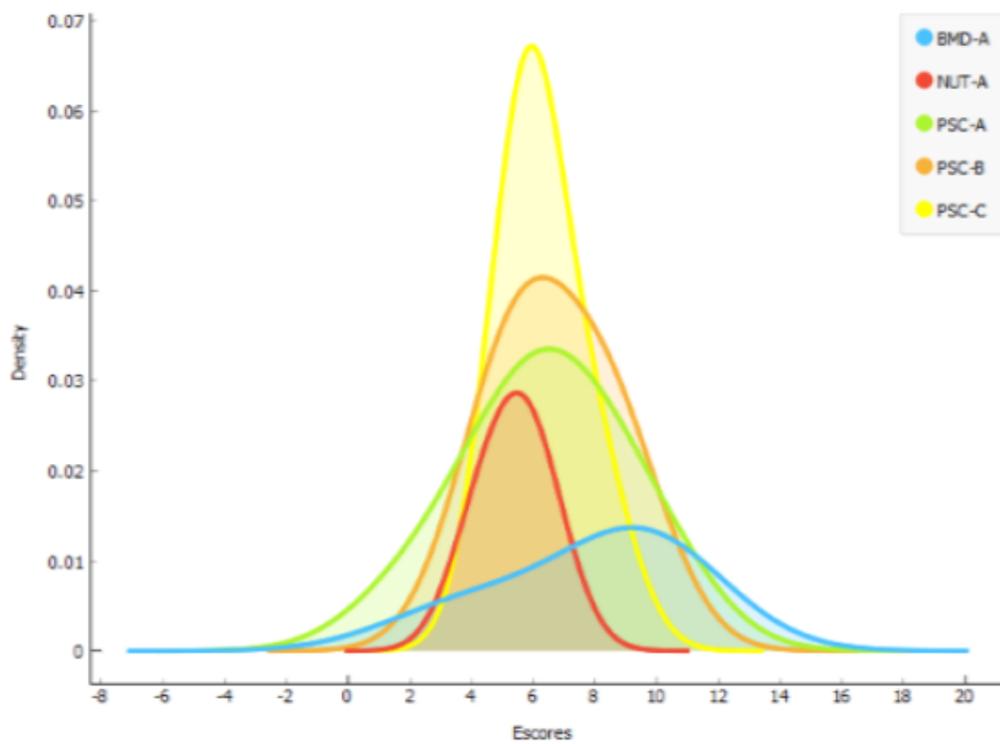


Figura 7.14: Curva de estimação da densidade de probabilidade núcleo gaussiano para os escores, separados por turma. (Gerado por Orange Canvas)

O teste de Shapiro-Wilk não permitiu assumir que todas as amostras provêm de distribuição normal, logo, a comparação entre as médias das turmas foi feita pelo teste de Kruskal-Wallis, que indicou que não há diferença significativa entre as médias das turmas ($p = 0,060$).

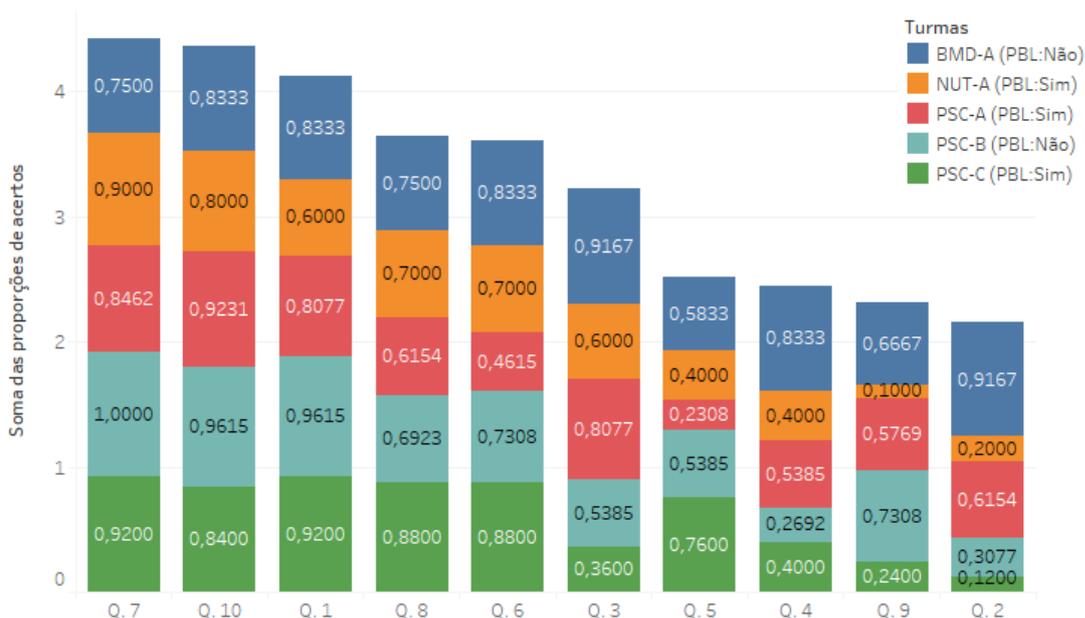


Figura 7.15: Proporções de acertos obtidos pelas diferentes turmas em cada questão (Gerado por Tableau).

A Figura 7.15 mostra que somente nas questões Q7 e Q4, a menor proporção de acertos alcançada não foi de uma turma que participou do PBL e nas questões Q7, Q8, Q6, Q5 a maior proporção de acertos foram de turmas que participaram do PBL.

A menor proporção de acertos entre todas as questões ocorreu na questão Q3 da turma PSC-C que participou do PBL.

Somente o curso de Psicologia teve turmas onde se pode comparar rendimentos estratificados pela participação no PBL e curso. Das três turmas de Psicologia, 51 (64,6%) alunos participaram do PBL (turmas PSC-A e PSC-C) e os 26 (32,9%) da turma PSC-B não participaram. Há uma grande semelhança entre as turmas e o teste de Kruskal-Wallis não apresentou diferença significativa entre os escores das turmas de Psicologia.

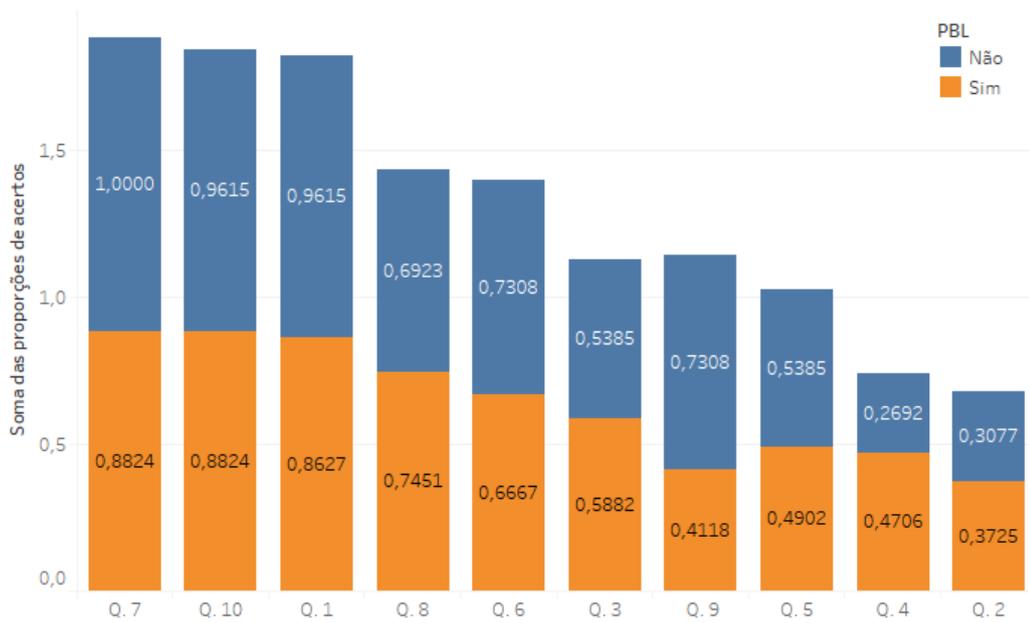


Figura 7.16: Proporções de acertos obtidos pelas turmas do curso de Psicologia separados por participação do PBL (Gerado por Tableau).

A maior diferença observada entre as proporções de acertos das turmas que participaram do PBL e as que não participaram ocorreu na questão Q9, em favor da turma que não participou do PBL.

As questões Q4 e Q2 foram as mais difíceis para as turmas de Psicologia, ao considerar-se a proporção de acertos. Nestas questões, a maior proporção de acertos foi das turmas que participaram do PBL.

Estratificação dos escores por participação na PBL

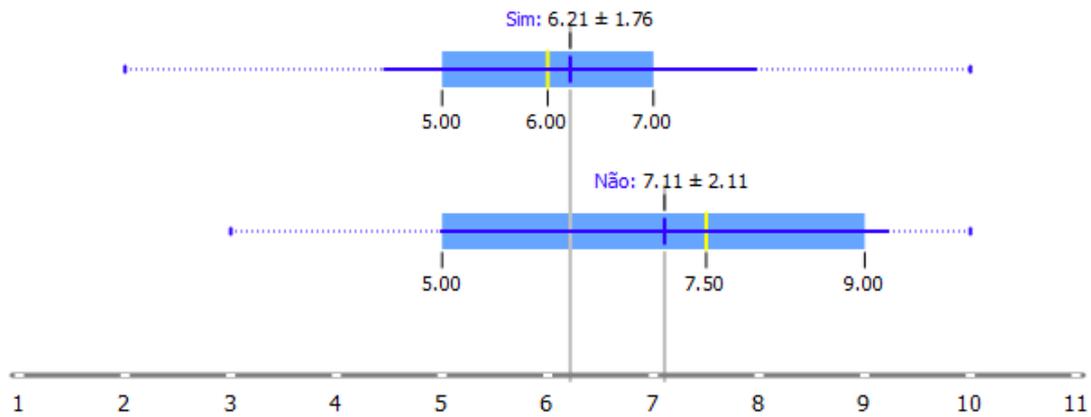


Figura 7.17: *Boxplots* de escore separado por participação no PBL. (Gerado por Orange Canvas)

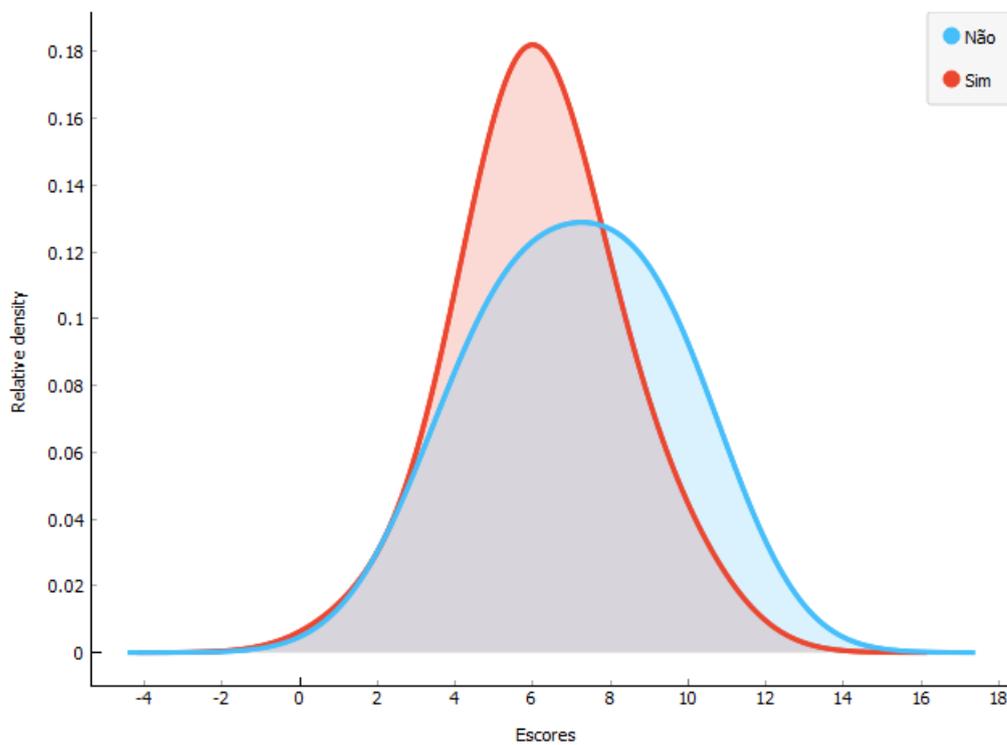


Figura 7.18: Curva de estimação da densidade de probabilidade núcleo gaussiano para os escores separados por participação no PBL. (Gerado por Orange Canvas)

Pela observação das Figuras 7.17 e 7.18, a média dos escores de quem participou da

PBL foi menor do que os demais, mas houve maior dispersão dos escores dos que não participaram.

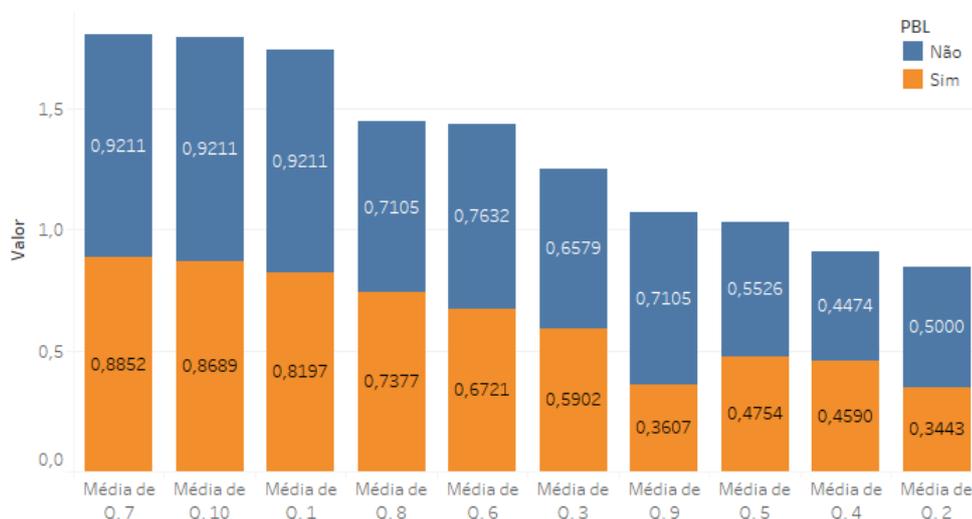


Figura 7.19: Distribuição dos escores e média dos escores separados por participação no PBL e por Q1, Q2, ..., Q10. (Gerado por Tableau)

Pela Figura ??, percebe-se que a proporção de acertos de quem não participou da PBL só não foi maior do que a dos que participaram nos itens Q8 e Q4 e, mesmo nestes itens, a diferença foi pequena. Isto aponta para uma consistência de melhor rendimento dos alunos “Não” para os alunos “Sim”.

O teste de normalidade de Shapiro-Wilk mostra que os escores não provêm de população normalmente distribuída quando se considera a participação ou não no PBL (os escores dos alunos que participaram do PBL obtiveram um $p = 0,02$, enquanto que os que não participaram tiveram um $p = 0,008$), indicando o uso de teste não paramétrico para a comparação dos dois grupos. Considerando o teste U de Mann-Whitney para amostras independentes, não rejeitamos a hipótese de igualdade entre os escores de quem participou do PBL e dos que não participaram da metodologia ($p = 0,059$).

Não há diferença significativa entre as médias dos escores quando estratificados pela variável PBL, embora as amostras indiquem que o escore médio dos que não

participaram da PBL supere a dos que participaram.

7.3 Nível de Domínio do CC dos THMP

O instrumento para a aferição do nível de domínio do CC dos THMP está definido na subseção 4.2.3, com as respostas esperadas às duas questões que foram aplicadas à amostra ao final do processo de PBL e levando em conta as sete categorias que englobam os invariantes operatórios necessários para o desenvolvimento satisfatório de tais questões, que são: Hipótese Nula, Hipótese Alternativa, Nível de Significância, Estatística do Teste, P-valor, Decisão e Interpretação.

Dos 99 indivíduos da amostra, apenas três não responderam às questões. As 96 respostas geraram dados analisados quanti e qualitativamente, como descrito a seguir.

7.3.1 Resultados Quantitativos

Para a análise quantitativa, cada uma das duas questões aplicadas aos 96 estudantes foram avaliadas segundo as sete categorias, de acordo com a pontuação 0, 1 ou 2 (descrita na subseção 4.2.3).

A tabela 7.20 mostra os valores descritivos para os escores totais. Nela, se vê que tanto a média quanto a mediana dos Alunos que não participaram do PBL foi maior do que a daqueles que participaram, tal como acontece com seus intervalos de confiança. A variabilidade, indicada por meio do Desvio Padrão, Intervalo e Amplitude interquartil, dos escores dos que participaram do PBL foi maior do que a dos que não participaram. Assim, os alunos que participaram do PBL tiveram uma média de escore total menor e uma dispersão maior do que os que não participaram. Entretanto, o teste não paramétrico de Mann-Whitney para amostras independentes não revelou diferença significativa² entre as distribuições do escore total separados por PBL ($p = 0,249^3$), ou

²Todos os testes consideram um nível de significância de 5%

³Todos os p-valores foram gerados no SPSS

seja as diferenças das médias não são significativas.

		Descritivos		
		PBL		Estatística
ESCORE TOTAL	Não	Média		13,29
		95% Intervalo de Confiança para Média	Limite inferior	11,23
			Limite superior	15,35
		Mediana		13,00
		Desvio Padrão		6,260
		Mínimo		0
		Máximo		23
		Intervalo		23
		Amplitude interquartil		11
	Sim	Média		11,24
		95% Intervalo de Confiança para Média	Limite inferior	9,23
			Limite superior	13,25
		Mediana		11,00
		Desvio Padrão		7,642
		Mínimo		0
		Máximo		24
		Intervalo		24
		Amplitude interquartil		12

Figura 7.20: Descritivo do escore total estratificado por PBL (Gerado no SPSS).

Os *boxplot* na figura 7.21 auxiliam na percepção das diferenças entre os escores dos dois grupos analisados, já apontadas em 7.20.

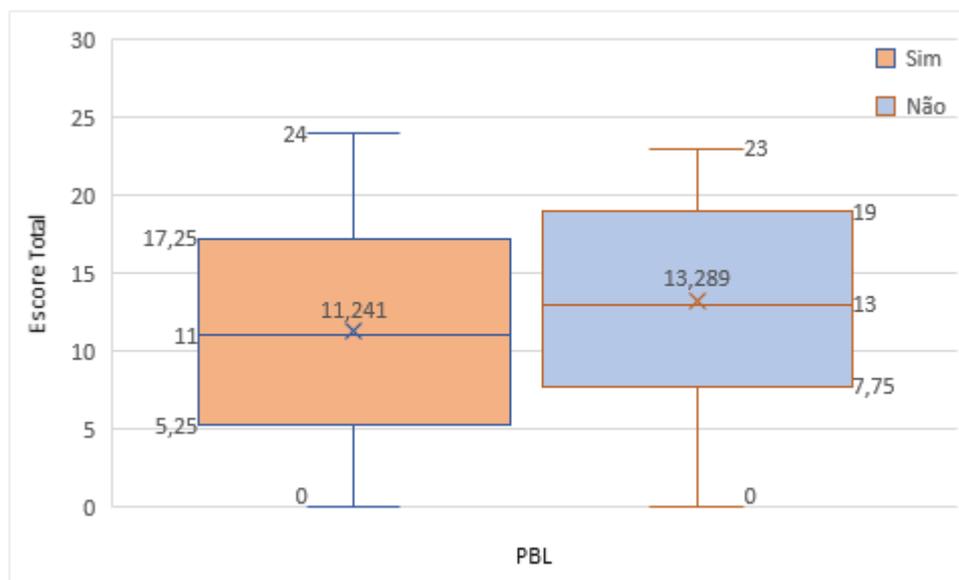


Figura 7.21: *Boxplots* comparativos do escore total estratificados por PBL (Gerado no Excel).

Vale lembrar que são sete categorias, avaliadas em duas questões, com pontuação mínima de zero e máxima de dois em cada uma das categorias, gerando uma pontuação mínima possível de zero e máxima de 28 pontos, onde se pode estabelecer os níveis de domínio do CC dos THMP, mostrados na tabela 7.4:

Tabela 7.4: Tabela de classificação do nível de domínio do CC dos THMP de acordo com o escore total das sete categorias do CC dos THMP em duas questões abertas.

Escore Total	Nível de Domínio
(0,10]	Baixo
(10,20]	Moderado
(20,28]	Ampla

A figura 7.22 mostra os histogramas dos escores separados pela participação ou não no PBL. Os valores percentuais referem-se à frequência relativa dos escores totais nos níveis de domínio indicados pela tabela 7.4.

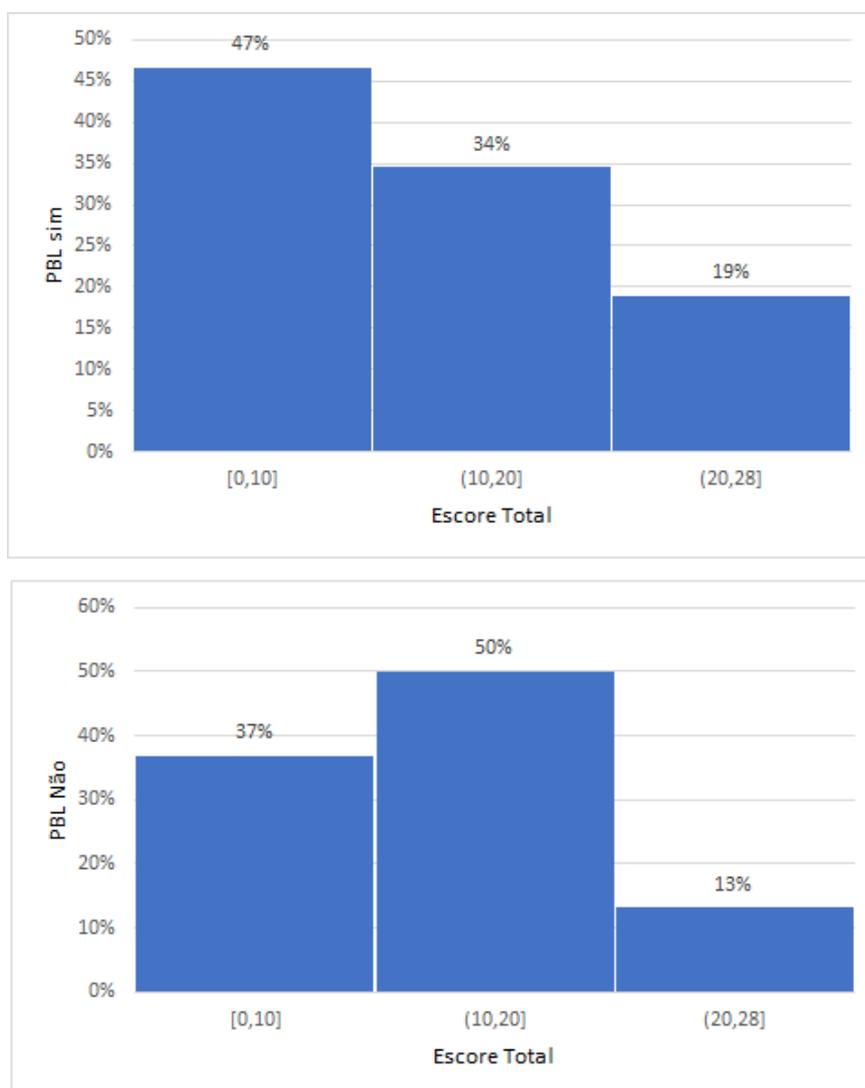


Figura 7.22: Distribuição de frequências do escore total dentro dos níveis de domínio para quem participou ou não do PBL (Excel).

Percebe-se que a metodologia tradicional (PBL não) aportou mais estudantes no nível moderado e menos no nível amplo, enquanto a metodologia do PBL teve uma maior incidência no nível baixo, porém, a frequência relativa no nível alto dos que participaram do PBL foi maior do que as dos que não participaram.

Focando a análise nos escores das questões (figuras 7.23 e 7.24), vê-se que a superioridade dos que não participaram do PBL se mantém quanto aos valores médios e aos

intervalos de confiança em torno das respectivas médias. As características da variabilidade também se preservaram em comparação aos escores totais. Novamente, como nos escores totais, nos escores das questões 1 e 2 não houve diferença significativa entre os que participaram e os que não participaram do PBL (p: 0,403 e 0,222, respectivamente, no mesmo teste utilizado para os escores totais.)

PBL		Estatística		
ESCORE(Q1)	Não	Média	7,16	
		95% Intervalo de Confiança para Média	Limite inferior	6,17
			Limite superior	8,14
		Mediana	7,00	
		Desvio Padrão	2,991	
		Mínimo	0	
		Máximo	12	
		Intervalo	12	
		Amplitude interquartil	4	
		Sim	Média	6,34
	95% Intervalo de Confiança para Média		Limite inferior	5,36
			Limite superior	7,33
	Mediana		7,00	
	Desvio Padrão		3,730	
	Mínimo		0	
	Máximo		13	
	Intervalo		13	
	Amplitude interquartil		5	

Figura 7.23: Descritivo do escore da questão 1 estratificado por PBL (SPSS).

ESCORE(Q2)	PBL		Estatística	
	Não			
	Não	Média		6,13
		95% Intervalo de Confiança para Média	Limite inferior	4,53
			Limite superior	7,73
		Mediana		7,50
		Desvio Padrão		4,861
		Mínimo		0
		Máximo		14
		Intervalo		14
		Amplitude interquartil		11
			Sim	Média
95% Intervalo de Confiança para Média	Limite inferior			3,42
	Limite superior			6,37
Mediana				3,00
Desvio Padrão				5,612
Mínimo				0
Máximo				14
Intervalo				14
Amplitude interquartil				10

Figura 7.24: Descritivo do escore da questão 2 estratificado por PBL (SPSS).

Para entender as diferenças detectadas em cada categoria de análise, a figura 7.25 mostra a pontuação média dos estudantes que participaram ou não do PBL, considerando a questão 1 (Q1), 2 (Q2) e ambas (Q1 + Q2). Lembrando que cada categoria dentro da questão um (Q1) ou dois (Q2) pode atingir um escore máximo de dois pontos implicando em quatro pontos possíveis para o escore total (Q1 + Q2) de cada categoria.

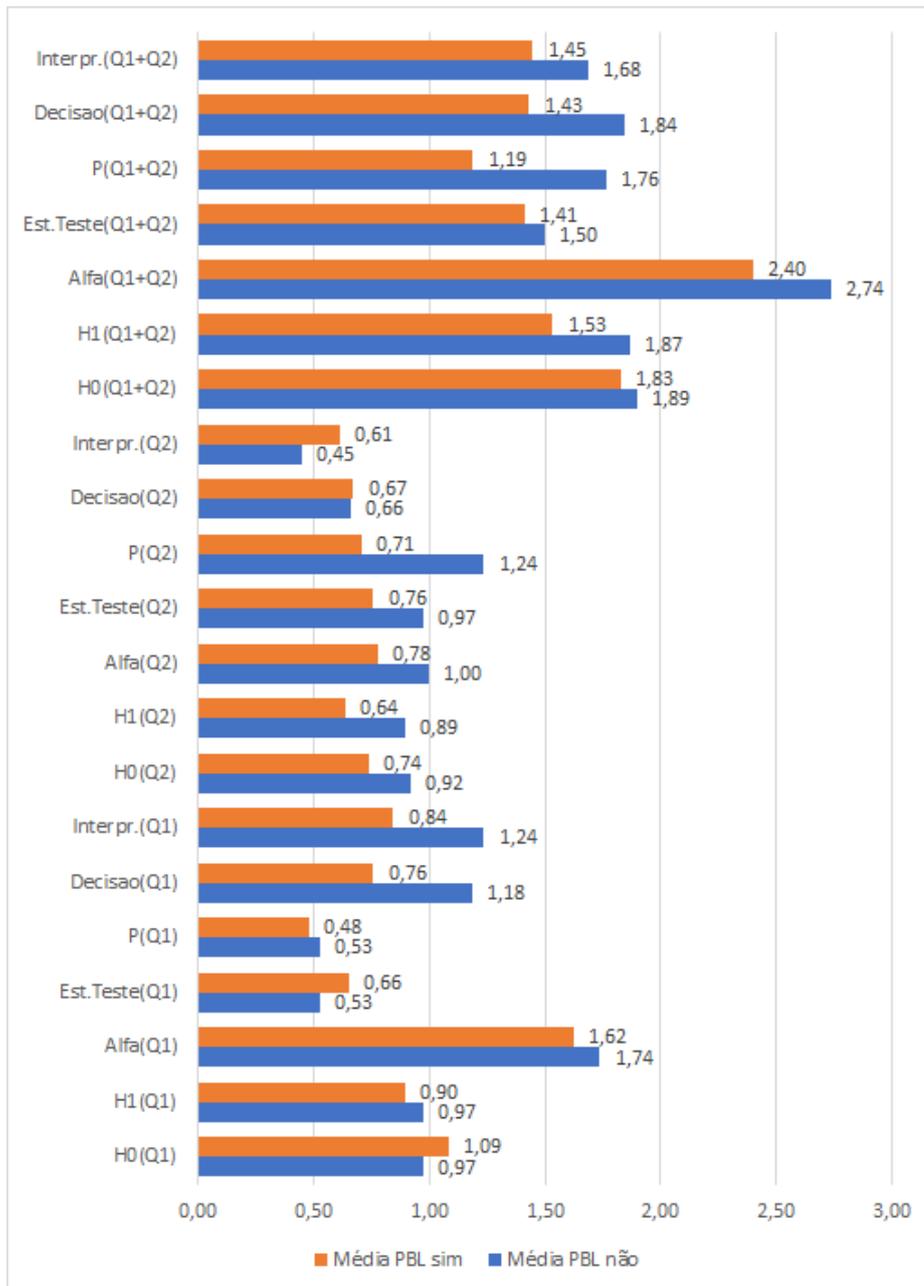


Figura 7.25: Médias das categorias estratificadas por PBL (Excel).

Somente as categorias H_0 (Q1), Estatística do Teste (Q1), Decisão (Q2) e Interpretação (Q2) tiveram uma média de pontuação melhor para os alunos que participaram da metodologia do PBL em relação aos que não participaram, sendo que a categoria

Decisão (Q2) teve valores praticamente iguais. Entretanto, o teste U de Mann-Whitney revelou que somente as categorias Decisão (Q1), Interpretação (Q1) e p-valor (Q2), possuem diferenças significativas entre as médias dos alunos que participaram do PBL e dos que não participaram, com médias maiores para os que não participaram da metodologia, nestas três categorias.

A fim de entender como as respostas se distribuíram nos diferentes níveis de domínio, em cada categoria, criou-se a figura 7.26. Nela, leva-se em conta as respostas às questões Q1 e Q2 tomadas juntas, o que implica na adição dos escores obtidos em Q1 com os da questão Q2, fazendo com que os escores originalmente dados entre zero e dois (descritos na seção 4.5), na figura 7.26, sejam tomados de zero a quatro, permitindo que se estabeleça a equivalência descrita na tabela 7.5.

Os valores nas colunas da figura 7.26, referem-se às quantidades percentuais de respostas em cada nível, dos 96 alunos, sendo 58 participantes do PBL (colunas vermelhas) e 38 não participantes (colunas azuis).

Tabela 7.5: Tabela de equivalência para os escores entre zero e dois (tabela 4.5) e zero e quatro (figura 4.5), para cada nível.

Nível	Escore de zero a dois	Escore de zero a quatro
Não apresenta/ Não operacional	0	0
Domina/Explicita Parcialmente	1	1 e 2
Domina/Explicita Totalmente	2	3 e 4

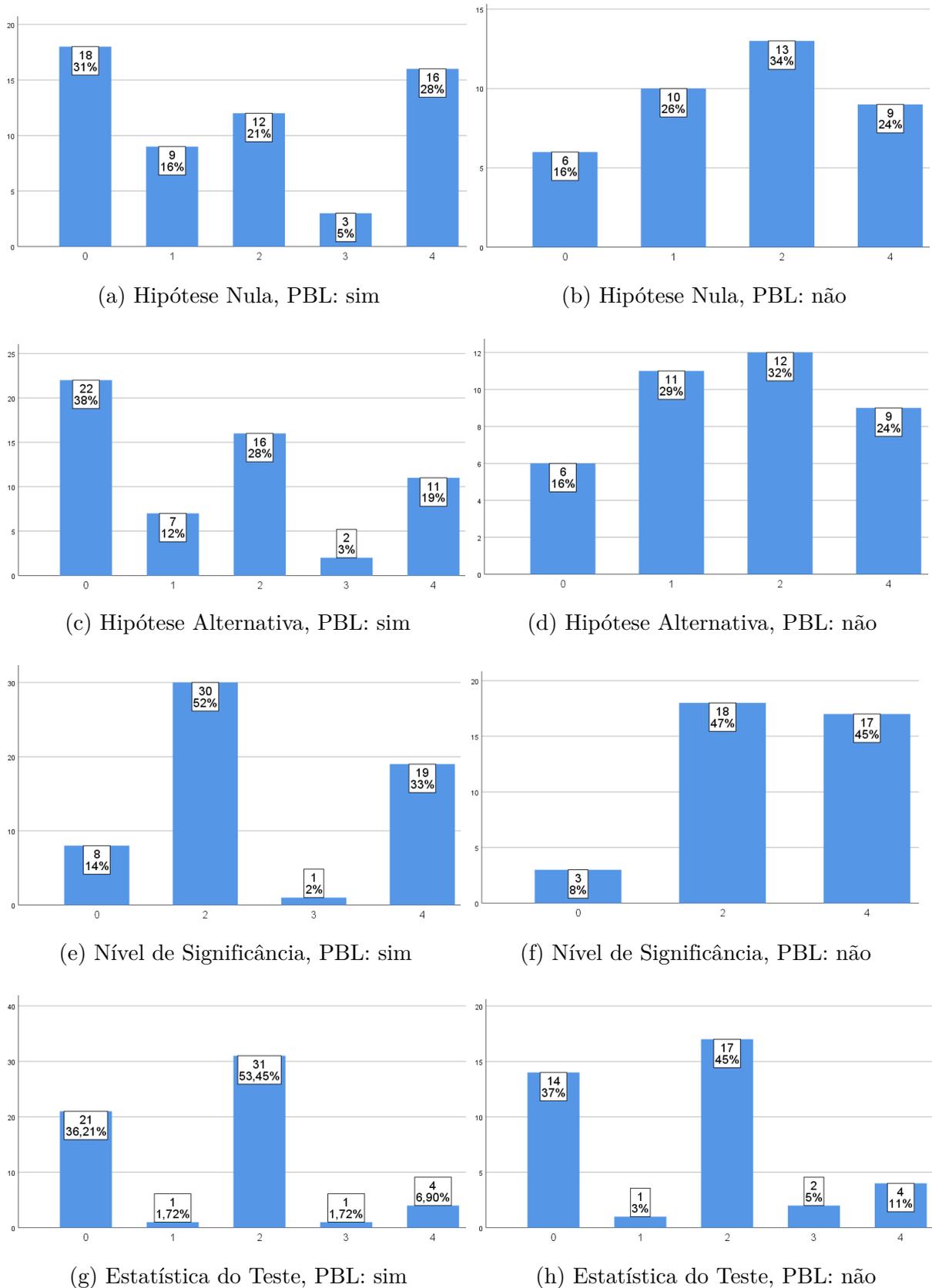


Figura 7.26: Distribuições dos alunos quanto ao nível atingido nas respostas das questões um e dois juntas, dentre as quatro primeiras categorias, separados pela participação no PBL.

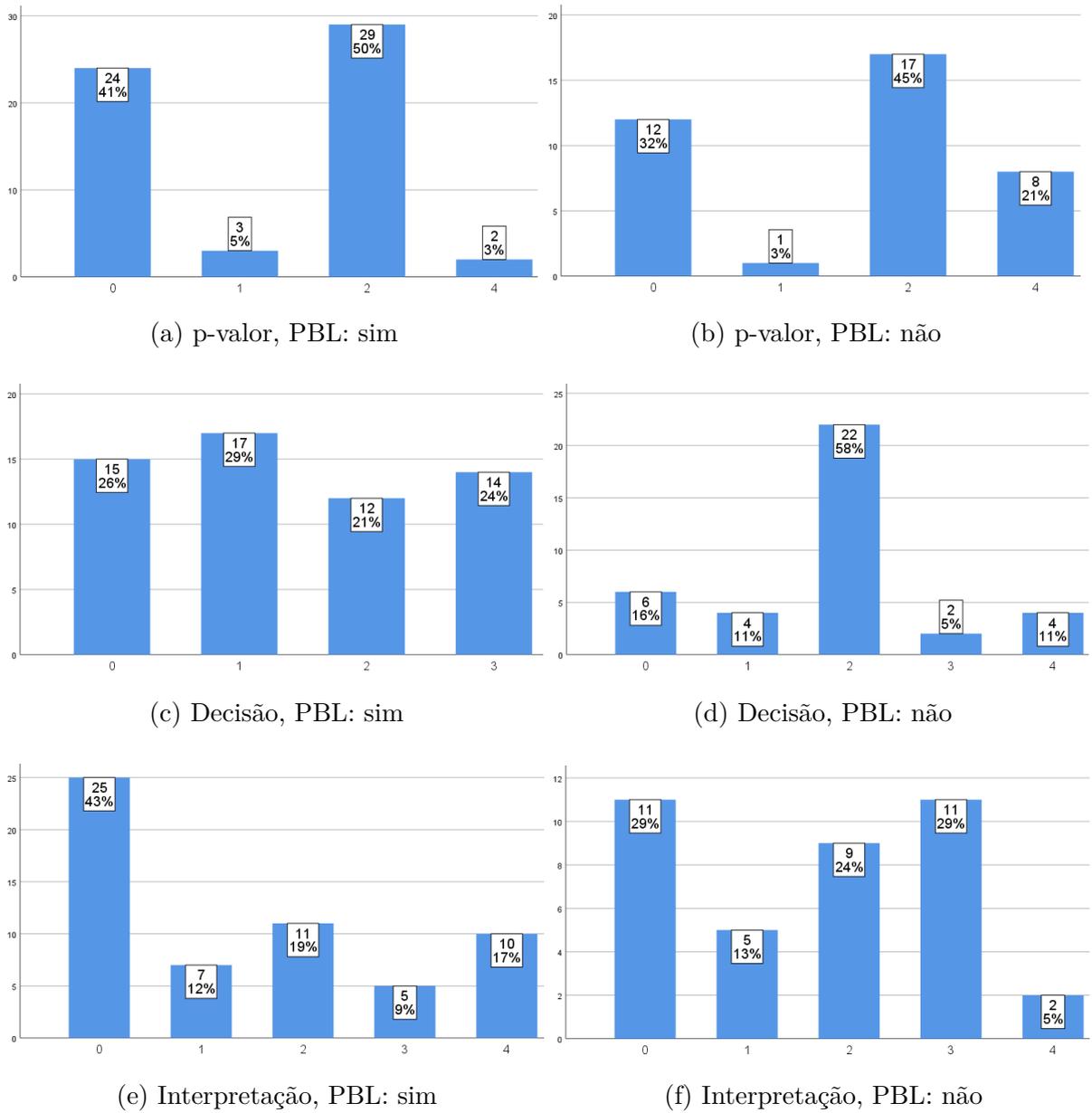


Figura 7.27: Distribuições dos alunos quanto ao nível atingido nas respostas das questões um e dois juntas, dentre as três últimas categorias, separados pela participação no PBL.

Por meio da figura 7.26, verifica-se que não há uma homogeneidade na distribuição das respostas em relação à Hipótese Nula (subfigura 7.26a) dos alunos participantes do PBL, visto que uma quantidade alta e semelhante destes alunos classificaram suas

respostas nos valores extremos da categoria (31% no nível mais baixo e 28% no mais alto). Os alunos não participantes do PBL tiveram um resultado mais uniforme ao longo das pontuações (7.26b), mas apenas nove deles mostraram domínio da categoria enquanto sua maior parte concentrou-se no nível intermediário de domínio da categoria.

Também na categoria Hipótese Alternativa, houve uma distribuição mais uniforme, nos diferentes níveis, dos alunos não participantes do PBL (subfigura 7.26d) do que nos participantes (subfigura 7.26c), porém, a quantidade de respostas que atingiram o maior nível de compreensão foi semelhante nos dois grupos (24% e 22%, respectivamente), o mesmo acontece com os níveis mais baixos, 45% para os não participantes e 50% para os participantes.

A categoria Nível de Significância teve percentuais altos para os níveis superiores nos dois grupos (subfiguras 7.26e e 7.26f), com uma e nenhuma resposta alocada no nível três, para os participantes do PBL e para os não participantes, respectivamente.

A categoria estatística do teste teve resultados muito parecidos em ambos os grupos (subfiguras 7.26g e 7.26h), concentrando as maiores frequências em um nível intermediário, seguidas do nível mais baixo de compreensão.

A distribuição para a categoria p-valor também foi semelhante em ambos os grupos (subfiguras 7.27a e 7.27b), com destaque para uma maior frequência no nível mais alto de compreensão para o grupo não participante do PBL.

Na categoria Decisão, já na figura 7.27, o grupo participante do PBL (subfigura 7.27c) teve uma distribuição mais uniforme por todos os níveis, a menos do nível quatro, enquanto que os não participantes (subfigura 7.27d) concentraram-se em um nível intermediário.

A categoria Interpretação teve grande incidência nos níveis mais baixos de compreensão, 55% no grupo participante do PBL (subfigura 7.27e) e 42% nos não participantes (subfigura 7.27f), enquanto que nos níveis mais altos sobressaíram-se as respostas do grupo não participante, mesmo que a maior frequência dos participantes no nível mais

alto tenha se destacado.

A análise quantitativa do instrumento para avaliação do nível de domínio do CC dos THMP vista até aqui, mostrou que as diferenças de rendimento encontradas não permitem descartar sumariamente o PBL, mas indica que outros estudos devem ser feitos para testar essa conjectura usando mais unidades experimentais. Os rendimentos levemente menores reforçam a ideia de que o esforço da adaptação do PBL para um contexto historicamente desenvolvido para adequar-se ao modelo tradicional pode gerar resultados, no mínimo, semelhantes. Outros ganhos esperados desta metodologia ativa de ensino são bem vindos, como a inserção do estudante em ambiente de pesquisa, o trabalho cooperativo, o protagonismo do aluno, a construção colaborativa do conhecimento, o autodidatismo, entre outras.

Findada a análise quantitativa do instrumento para o estabelecimento do nível de domínio do CC dos THMP, passa-se à análise pormenorizada das duas questões que subsidiaram a análise quantitativa. Neste sentido, a análise qualitativa procura indícios de invariantes operatórios que por ventura tenham escapado das categorias elencadas ou por esquemas de pensamentos desenvolvidos pelos estudantes que permitam um entendimento melhor de quão bem os alunos manejam este Campo Conceitual.

7.3.2 Resultados Qualitativos

O instrumento para medir o nível de domínio aparente do CC dos THMP possui sete categorias de análise para as duas questões que foram aplicadas. Os desenvolvimentos esperados em cada categoria, para as duas questões, estão descritos na subseção 4.2.3 e foram levados em conta nos resultados da subseção anterior (7.3.1). As análises qualitativas desta subseção referem-se ao mesmo instrumento e complementam as análises quantitativas da subseção anterior. Neste sentido, estas análises qualitativas não são apresentadas de forma extensiva, pois não são a única análise do instrumento e demandariam um espaço demasiadamente grande da Tese.

Após a leitura detalhada de todas as respostas dos alunos que participaram do PBL, procedeu-se à escolha daquelas capazes de representar uma boa resposta a cada questão, bem como uma resposta ruim a ambas.

Relembra-se as questões contidas no instrumento e apresenta-se as respostas dos alunos, a seguir:

1. A taxa de fenilalanina no soro, em indivíduos normais, é de 1,4 mg/100mL e o desvio padrão é de 0,32 mg/100mL. Supõe-se que cem pacientes formam uma amostra aleatória da população de indivíduos normais, com média amostral de 1,3 mg/100mL.

Considerando um nível de significância de 1%, os dados estão de acordo com a suposição feita?

1) $H_0: \mu = 1,4$
 $H_1: \mu \neq 1,4$
 ↳ bilateral porque é uma amostra aleatória, então é diferente se encicla melhor no que precisa mais

2) $\alpha = 0,01$ (dados fornecidos na questão) 2

3) $p = 0,0057$ 2E
 2P

4) $p < \alpha \rightarrow$ rejeita H_0
 $p > \alpha \rightarrow$ não rejeita H_0

$0,0057 < 0,01 \rightarrow$ rejeita o H_0 1

Considerando o nível de significância de 1%, os dados experimentais não estão de acordo com a suposição feita, pois o p-valor (probabilidade de significância) é menor que α (Nível de significância).
 O teste escolhido no bimestre, é o teste z pois o valor de n (número da amostra) é superior a 30.
 Portanto a suposição feita não satisfaz nossas expectativas. 2

Figura 7.28: Uma boa resposta à questão 1, dada pelo estudante ID25 . Os grifos coloridos são anotações do professor/tutor.

O aluno inicia escolhendo adequadamente as hipóteses nula e alternativas, mas justifica erroneamente o teste bilateral por se tratar de uma amostra aleatória.

Nota-se que os quatro passos da estruturação do conteúdo dos THMP estão bem destacados e, no item 3 (cálculo do p-valor) percebe-se que o processo de cálculo ocorrido no *software* Bioestat e a escolha do p-valor, aconteceram a contento, com a justificativa de escolha do teste feita pelo tamanho amostral ser maior do que 30 e não levando em conta a posse do desvio padrão. Cabe lembrar que os alunos estavam no laboratório de informática.

No quarto passo, o aluno expõe os critérios de decisão antes de comparar o p-valor com o nível de significância e acaba decidindo corretamente a rejeitar a hipótese nula.

A interpretação do resultado afirma que os dois dados, supostamente, média e desvio padrão, não estão de acordo com a suposição feita. Parece que o aluno está analisando os valores da média e do desvio padrão de forma desconexa com a amostra subjacente a eles, como se estes dois resultados fossem a representação total da amostra. Este parece ser um indicador, observado em muitas outras respostas, de que as questões sobre os processos amostrais concernentes à variabilidade, não estão bem compreendidas.

Por fim, percebe-se que a parte mecânica (algorítmica), e de tomada de decisão foram realizadas pelo estudante, mas a interpretação sofre com imprecisões.

Passa-se a analisar a resposta mostrada na figura [7.29](#), que apresenta problemas conceituais bastante relevantes.

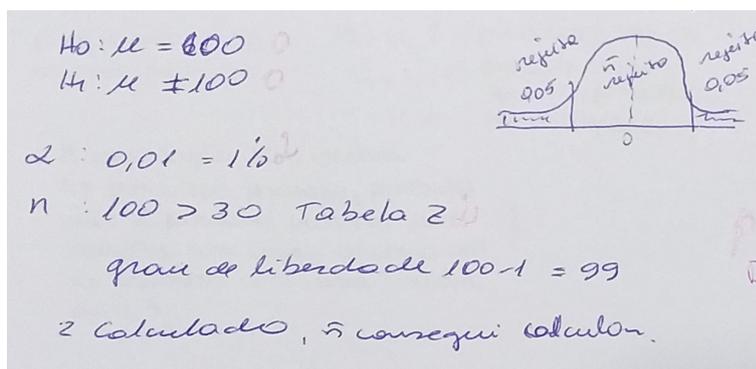


Figura 7.29: Uma resposta com problemas conceituais à questão 1, dada pelo estudante ID86. Os grifos coloridos são anotações do professor/tutor.

Os problemas conceituais mostram um entrelaçamento das abordagens do p-valor e da região de rejeição da hipótese nula (expostas na subseção 3.3.2), ambos incorretos no contexto apresentado.

O aluno percebe que o processo inicia com o estabelecimento das hipóteses nula e alternativa mas o faz usando o tamanho da amostra e não o valor de referência. O tamanho amostral é usado corretamente para a determinação do teste a utilizar e do grau de liberdade, mas o processo não alcançou o estabelecimento da estatística do teste nem da zona de rejeição.

Aparentemente, o aluno não conseguiu calcular a estatística do teste por não dominar o *software* Bioestat e a zona de rejeição, por não dominar o uso da tabela Z. Assim, a incompletude da resposta parece recorrer da incapacidade de reproduzir o algoritmo que leva aos números faltantes e, neste caso, o algoritmo parece ter inviabilizado a manifestação do conceito.

Agora, passa-se a duas repostas à questão dois.

2. A média de acidentes mortais em uma cidade é de 12 mensais. Após uma campanha de conscientização sobre os perigos no trânsito, nos seis meses consecutivos contabilizaram-se 8, 11, 9, 7, 10 e 9 acidentes mortais. A campanha foi efetiva?

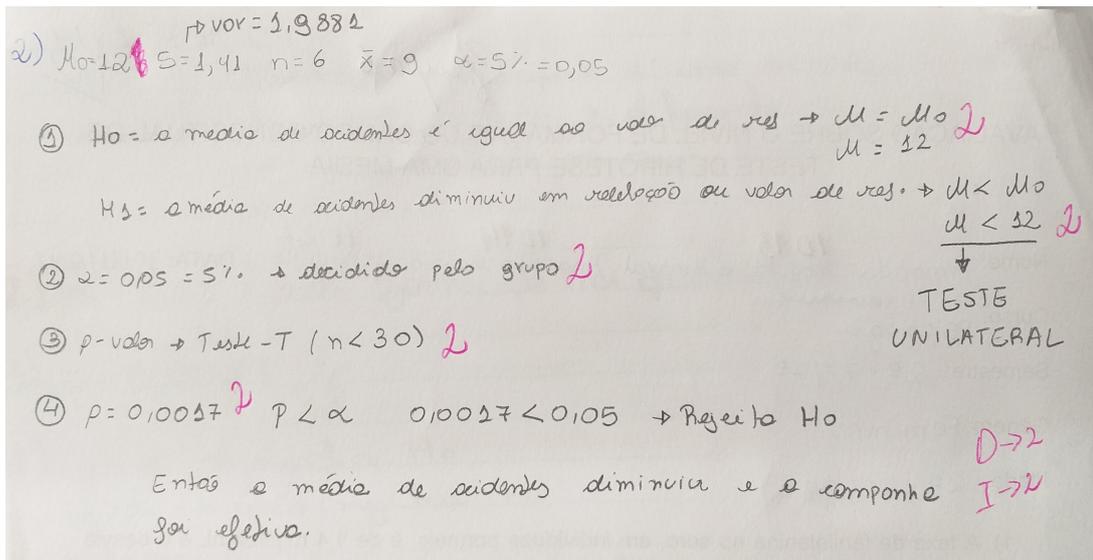


Figura 7.30: Uma resposta adequada à questão 2 dada pelo estudante ID22 . Os grifos coloridos são anotações do professor/tutor.

O aluno inicia a resolução calculando as medidas amostrais da média, variância e desvio padrão e parte para o THMP propriamente dito, o que foi feito preservando a estrutura dos THMP.

Em seu item 1, o aluno define as hipóteses nula e alternativa usando duas representações, a escrita cursiva e a notação matemática própria dos THMP. Este fato corrobora com a existência de um conceito em ação bem estabelecido, que permite o reconhecimento do conceito de hipóteses estatísticas e é capaz de explicitar-se de mais de uma maneira. Além disso, o entendimento do problema parece ter conduzido o aluno a escolher a opção $\mu < 12$ para a hipóteses nula, pois a escolha “algorítmica” seria $\mu \neq 12$ para a hipótese nula.

A determinação do nível de significância, por sua vez, trouxe a explicação do por quê o aluno escolheu 5%, atribuindo a escolha ao grupo a que ele pertencia durante o processo do PBL, como se a escolha feita anteriormente servisse para todos problemas sobre THMP, tal como muitos pesquisadores fazem, tendo por grupo, seus pares da área da pesquisa.

O cálculo do p-valor, na etapa três, foi bem conduzido no *software* Bioestat, dado o valor correto e as indicações do teste utilizado. Novamente, como o aluno ID25 (figura 7.28), não houve menção sobre a variância ou desvio padrão para a escolha do teste, deixando implícita o uso destas medidas para tal escolha.

A tomada de decisão pela comparação do p-valor com o nível de significância foi correta, entretanto, a interpretação subsequente levou em conta, novamente, a média amostral, e não o processo amostral que fundamenta o teste, reiterando a existência de um possível invariante operatório incorreto: “Ao rejeitar-se a hipótese nula, a média amostral corresponde à hipótese alternativa”.

A próxima resposta, mostrada na figura 7.31, traz problemas conceituais relativos à questão dois.

$\text{média} = 12 / \text{mês}$
 campanha \rightarrow 6 meses depois
 1° 2° 3° 4° 5° 6°
 $\frac{8 + 11 + 9 + 7 + 10 + 9}{6} \rightarrow \text{média} = 9$
 $\text{desv. padrão} = 1,4$
 $p = 0,035$
 $H_0 = 12$
 A campanha não foi repetida

Figura 7.31: Uma resposta inadequada à questão 2 dada pelo estudante ID82. Os grifos coloridos são anotações do professor/tutor.

O aluno parece ter usado os valores da média amostral e do desvio padrão para o cálculo do p-valor e para a tomada de decisão, mas como não explicitou minimamente este processo, não se pode extrair maiores inferências a respeito. Este raciocínio excessivamente implícito pode esconder a falta de conhecimento mais consistente acerca do tema ou a não tão rara sensação de que somente expondo valores responde-se a uma questão de Estatística, como é comum em questões de matemática do ensino médio, por exemplo.

Abaixo, segue uma análise das respostas atípicas (inesperadas ou não previstas na análise quantitativa (7.3.1) em cada categoria, para as duas questões. Esta análise é auxiliar na ampliação da compreensão do domínio do CC dos THMP, alcançado pelos estudantes da amostra. Assim, espera-se encontrar indícios de invariantes operatórios que estejam além dos esperados e já avaliados nas sete categorias anteriormente definidas.

Hipótese Nula Não raro, os alunos escolheram escrever a hipótese nula como $H_0 = 1,4$ na questão 1 (ou $H_0 = 12$, na 2). Este tipo de escrita dificulta a identificação do conceito em ação subjacente, pois explicitamente remete à “a hipótese nula é 1,4”, que é incorreto, mas implicitamente, pode remeter à “o valor de referência (média populacional) a ser usado na hipótese nula é 1,4”, que está correto. A sequência dos cálculos nem sempre foi capaz de elucidar esta questão, não permitindo inferir se o conceito em ação correto $H_0 : \mu = 1,4$, era a manifestação implícita para $H_0 = 1,4$.

Um nível de interpretação diferenciado pode ser visto na figura 7.32 onde o aluno refere-se à hipótese nula sem usar a notação típica, mostrando um bom entendimento da função da hipótese nula, com a explicitação do invariante operatório correto por meio de expressão não matemática.

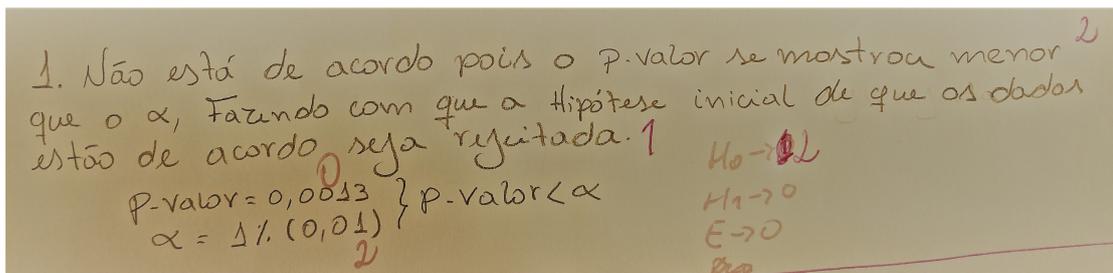
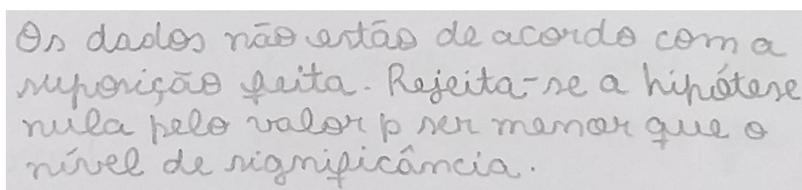


Figura 7.32: Resposta do estudante ID16 à questão 1. Os grifos coloridos são anotações do professor/tutor.

Hipótese Alternativa A hipótese alternativa é também chamada de hipótese do investigador, dada a sua relevância principalmente quando da rejeição da hipótese

nula. Mesmo assim, nem sempre a hipótese alternativa foi explícita nas questões, como a figura 7.33 exemplifica. Este fato pode indicar uma fragilidade no entendimento do papel da hipótese alternativa para além da decisão do teste uni ou bilateral. Neste sentido, o conceito em ação que identifica a hipótese alternativa como substituta à hipótese nula em caso de rejeição desta, nem sempre está presente.

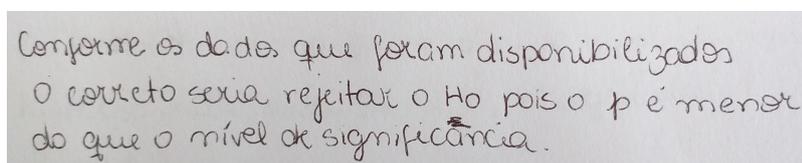


Os dados não estão de acordo com a suposição feita. Rejeita-se a hipótese nula pelo valor p ser menor que o nível de significância.

Figura 7.33: Resposta do estudante ID79 à questão 1.

Nível de Significância Apareceram somente dois níveis de significância: 1 e 5%, com larga diferença em favor dos 5%. Não houve explicação sobre o motivo da escolha, que figurou como um valor de referência para a tomada de decisão (em geral, para comparar com o p-valor), como exemplifica a figura 7.34. Parece ser comum o conceito em ação incorreto “o nível de significância é 5%”. Pareceu que o uso do nível de significância foi entendido somente em seu nível operacional, sem a preocupação do entendimento do seu significado mais abrangente.

Por outro lado, não houve menção ao nível de confiança o que pode sugerir que a diferença com o nível de significância esteja clara.



Compare os dados que foram disponibilizados. O correto seria rejeitar o H_0 pois o p é menor do que o nível de significância.

Figura 7.34: Resposta do estudante ID87 à questão 2.

Estatística do Teste (Z ou T) O *software* que os alunos dispunham (Bioestat) exige

que se escolha a estatística a ser utilizada (alguns softwares tem este processo automatizado) antes de mostrarem o valor da estatística e do p-valor, o que obriga o aluno a tomar uma decisão que nem sempre apareceu explícita na resolução do exercício, como no desenvolvimento mostrado na figura 7.35. Nela, se percebe que o aluno escolheu a distribuição T e o p bilateral por conta do resultado apresentado ($p = 0,0023$), que está incorreto. Além disso, o software pede pela variância da amostra, exigindo que o aluno eleve o desvio padrão quadrado, conhecimento também implícito neste caso e que foi feito corretamente.

Dado que muitos alunos escolheram esta resolução, pode estar disseminado o conceito em ação equivocado “para THMP usa-se teste t”.

$$\begin{aligned}
 H_0 &= \mu \neq 1,4 \\
 H_1 &= \mu \neq 1,4 \\
 t &= 11,2 \\
 P_{\text{valor}} &= 0,0023 \\
 P < t &\rightarrow \text{rejeita } H_0
 \end{aligned}$$

Figura 7.35: Desenvolvimento do estudante ID90 à questão 1. Os grifos coloridos são anotações do professor/tutor.

P-valor ou Estatística z ou t Considerando que o processo de cálculo foi automatizado pelo *software* Bioestat, os problemas detectados nesta categoria foram acerca da escolha da distribuição do teste (Z ou T) ou do tipo de teste (uni ou bilateral) e não ao processo de cálculo em si.

Decisão Chama atenção o fato da tomada de decisão aparecer muitas vezes sem considerar a hipótese alternativa, o que parece fragilizar o processo decisório visto que nem sempre é possível perceber a aceitação da hipótese alternativa pela resolução do aluno, como na resolução da figura 7.36.

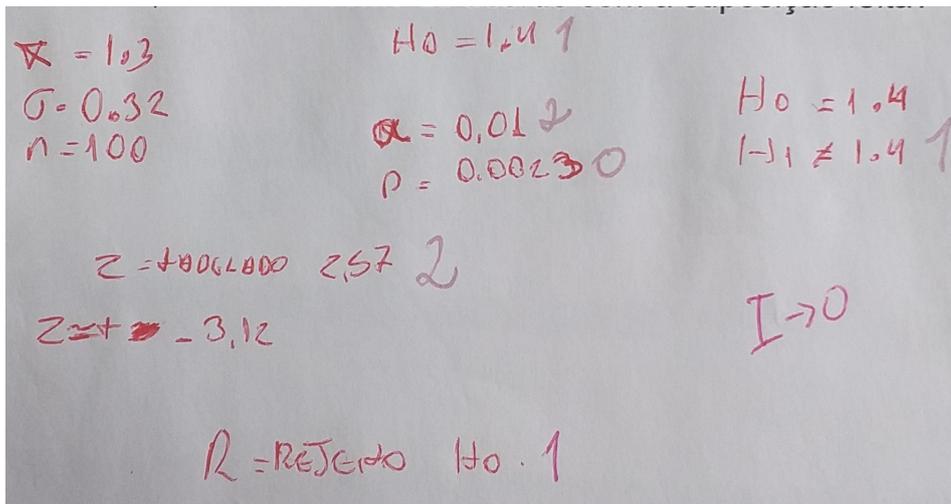


Figura 7.36: Desenvolvimento do estudante ID39 à questão 1. Os grifos em vermelho mais claro são anotações do professor/tutor.

Não raro, o estudante finaliza o processo dos THMP nesta fase, excluindo a contextualização e interpretação do seu resultado, como a resolução da figura 7.37, onde há uma preocupação com a parte mecânica do processo em detrimento da sua interpretação.

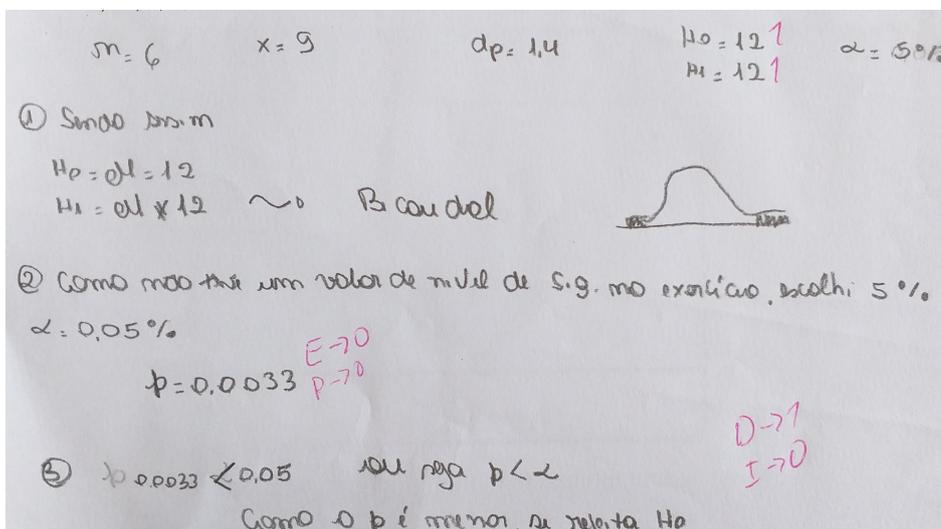


Figura 7.37: Desenvolvimento do estudante ID95 à questão 1. Os grifos em vermelho mais claro são anotações do professor/tutor.

Interpretação Nem sempre as interpretações foram provenientes do THMP e, por vezes, conceitos anteriores aos THMP foram elencados de maneira inadequada ou incompleta, como a interpretação mostrada na figura 7.38. Mesmo que o esforço do uso de conceitos estatísticos prévios seja louvável, a confusão entre eles mostra que não houve apropriação destes conceitos, como a interpretação mostrada na figura 7.39, que confunde o estabelecimento da hipótese nula com a tomada de decisão do teste.

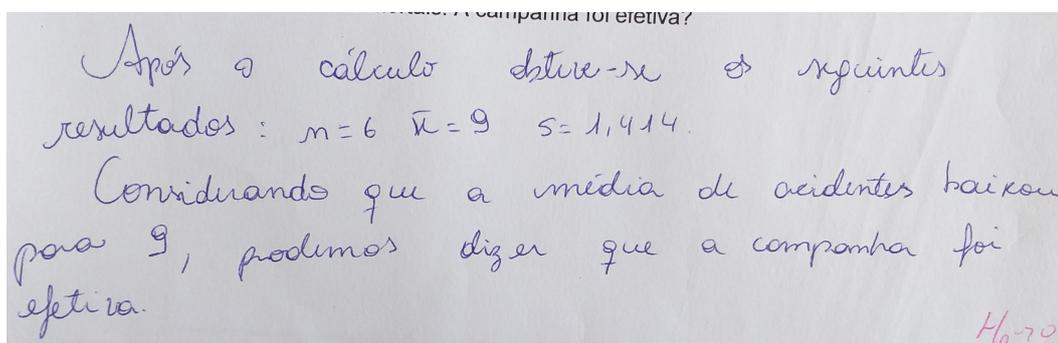


Figura 7.38: Desenvolvimento do estudante ID81 à questão 2. O grifo em vermelho é do professor/tutor.

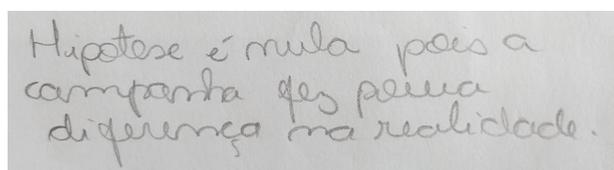


Figura 7.39: Interpretação do estudante ID19 à questão 2.

Também apareceram respostas atípicas fora das categorias de análise, como as seguintes:

- É esperado que a decisão do teste e sua interpretação refiram-se às hipóteses nula e alternativa por meio de sua notação usual. Não deixa de surpreender, respostas que fogem desta notação além de colocar uma abordagem menos comum, mas correta, dos THMP. Isto pode ser visto na resposta do estudante ID 76, que, no

lugar de simplesmente rejeitar ou não a uma hipótese em benefício da outra, usou a abordagem da plausibilidade da obtenção da amostra (figura 7.40).

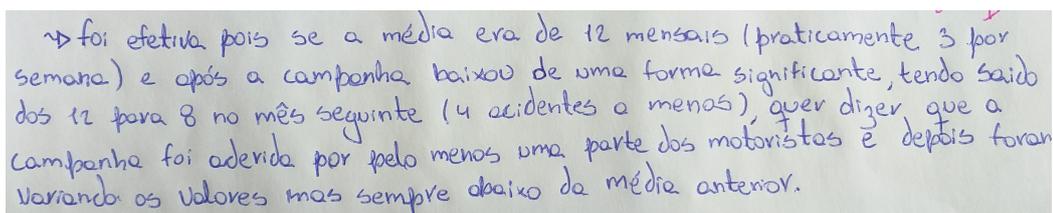
Concluiu que:
A suposição inicial foi de que a média era de 1.4 mas determinamos que não é plausível retirar uma amostra de 100 pessoas com média 1.3, considerando um alfa de 1%.

Figura 7.40: Resposta do estudante ID76 à questão 1.

- Outra abordagem foi a análise somente da média da amostra na questão 2, suscitando uma abordagem alheia aos conceitos de variabilidade contidos na questão, como a interpretação do estudante da figura 7.41. Neste sentido, o uso de termos estatísticos de forma equivocada, como o uso da palavra “significante” vista na resolução da figura 7.42, denota a fragilidade de uma interpretação baseada somente na análise descritiva da amostra. Esta fragilidade no pensamento inferencial também pode ser observada na interpretação do aluno da figura 7.43.

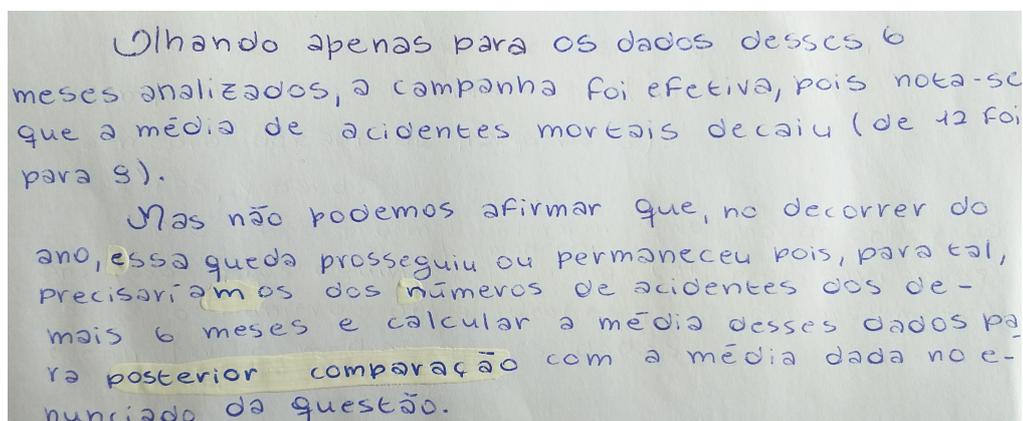
Pode-se afirmar que a campanha não foi efetiva pois a média de acidentes mortais diminuiu segundo a amostra. Pois antes era de 12 mortes por mês, e após a campanha a média foi de 9 mortes por mês.

Figura 7.41: Resposta do estudante ID84 à questão 2. Os grifos em vermelho são anotações do professor/tutor.



↳ foi efetiva pois se a média era de 12 mensais (praticamente 3 por semana) e após a campanha baixou de uma forma significativa, tendo saído dos 12 para 8 no mês seguinte (4 acidentes a menos), quer dizer que a campanha foi aderida por pelo menos uma parte dos motoristas e depois foram variando os valores mas sempre abaixo da média anterior.

Figura 7.42: Resposta do estudante ID43 à questão 2.



Olhando apenas para os dados desses 6 meses analisados, a campanha foi efetiva, pois nota-se que a média de acidentes mortais decaiu (de 12 foi para 9).

Mas não podemos afirmar que, no decorrer do ano, essa queda prosseguiu ou permaneceu pois, para tal, precisaríamos dos números de acidentes dos demais 6 meses e calcular a média desses dados para posterior comparação com a média dada no enunciado da questão.

Figura 7.43: Resposta do estudante ID45 à questão 2.

- Chama atenção uma resposta totalmente discursiva que remete às questões cruciais dos THMP sem deixar de abordar a análise descritiva da amostra, transmitindo a ideia de que o aluno transita com boa desenvoltura na maioria dos conceitos relevantes ao tema. Sua resposta à questão dois pode ser vista na figura 7.44.

foi constatado que os 6 meses seguintes apresentam uma média de acidentes mortais de 9 e desvio padrão de 1,4.

Considerando 12 como valor de referência e a hipótese alternativa sendo menor que 12 e o nível de significância sendo 19%.²

se chegou ao p-valor de 0,17%, sendo menor que o nível de significância, rejeitando o valor de referência e aceitando a hipótese de que a média de acidentes mortais seja menor que 12.

Dessa forma, os resultados apresentados contribuem para a afirmação de que a campanha foi efetiva.

$H_0 \rightarrow 2$

Figura 7.44: Resposta do estudante ID55 à questão 2. Os grifos em vermelho são do professor/tutor.

Alguns alunos chamaram atenção pelo seus desenvolvimentos das questões:

7.3.3 Resumo dos Resultados

As análises quantitativa e qualitativa mostraram um baixo nível de domínio do CC dos THMP tanto para o grupo que passou pelo PBL quanto para os que tiveram aulas tradicionais, efeito, acredita-se, ocasionado por questões estruturais e pessoais.

Estruturalmente, a disciplina, ao contrário de sua relevância, sofre com uma carga horária muito baixa nos currículos dos três cursos, tornando processos que evidenciam a pesquisa e a atividade do aluno, costumeiramente mais lentos, muito difíceis de serem implementados. Também há pouca valorização da disciplina no cursos analisados indo de encontro à tendência de crescimento da saúde baseada em evidências.

Em termos pessoais, os alunos demonstraram um grau de ansiedade estatística elevado, muito motivado pela relação com a matemática. Foi importante, também, a dificuldade de adaptação a uma metodologia que exige mais atividade do aluno, onde o estudar para a prova os conteúdos que o professor apresentou, já não é suficiente. Este ponto aparece com alta relevância quando se pondera que os instrumentos quantitativos adotados neste trabalho, basearam-se em questões típicas do modelo tradicional de

ensino, desenvolvido, acredita-se, ao longo de toda a formação dos estudantes, enquanto que a metodologia ativa do PBL, restringiu-se a poucas semanas.

Houve, também, um problema que parece compartilhar sua origem em questões estruturais e pessoais ao mesmo tempo, que é a carência formativa, especialmente em matemática, do aluno ingressante nos cursos analisados. Não seria demasiado achar que alunos que escolhem a área da saúde, principalmente o curso de Psicologia, tendem a não gostarem de matemática e ao que julgam ser derivados, como a estatística.

Outros motivos para os resultados obtidos estão expostos no capítulo da discussão (capítulo 8). Por hora, a tabela 7.6⁴ traz um resumo comparativo dos resultados quantitativos e qualitativos acerca do instrumento para medição do domínio do CC dos THMP.

⁴Para diminuir o tamanho do texto, nesta tabela, os alunos participantes do PBL são chamados de “PBL sim” enquanto que os não participantes, de “PBL não”.

Tabela 7.6: Comparação dos resultados quantitativos e qualitativos para o instrumento para determinação do nível de domínio do CC dos THMP, por categoria.

Categoria	Resultado Quantitativo	Resultado Qualitativo	Conclusão
Hipótese Nula	Ambos os grupos com rendimentos intermediários, mas com distribuições diferentes ao longo dos níveis, com o grupo sim posicionando mais respostas nos níveis mais altos.	Diferentes representações do conceito nem sempre deixam claro se os alunos o compreenderam, dificultando a identificação do conceito em ação.	Não raro há problemas de identificação da hipótese nula como hipótese a ser testada, bem como na identificação de qual deve ser o valor de referência populacional dentro da situação.
Hipótese Alternativa	Rendimentos semelhantes ao da categoria hipótese nula, em ambos os grupos.	Não raro esteve implícita nas respostas dos alunos. Esta categoria e a anterior, nem sempre figuram como de explicitação essencial pelos alunos.	A necessidade do estabelecimento da hipótese alternativa logo após a hipótese nula nem sempre surgiu na etapa um, tornando-se necessária durante a escolha do p-valor ou na rejeição da hipótese nula.
Nível de Significância	Os grupos posicionaram fortemente nos níveis mais altos, sem diferenças expressivas entre os grupos.	Estabelecido em 5% <i>a priori</i> , como se fosse um valor intrínseco ao uso dos THMP.	Esta categoria figurou como algo pré-estabelecido, definido pela situação ou tomado o valor de 5% como padrão. Não houve aprofundamento neste conceito.
Estatística do Teste	Grupos apresentaram um nível intermediário de compreensão, porém, houve uma frequência importante de respostas posicionadas no nível zero, em ambos os grupos.	A escolha do teste ficou implícita no processo de manuseio do <i>software</i> Biostat.	A categoria figurou de maneira processual, sem menção a conceitos mais profundos, como as distribuições amostrais.

Tabela 7.7: Continuação da tabela 7.6.

Categoria	Resultado Quantitativo	Resultado Qualitativo	Conclusão
p-valor	Distribuição muito semelhante à categoria anterior, com ambos os grupos concentrando suas respostas no nível intermediário e razoável frequência no nível zero.	Se a categoria anterior foi bem desenvolvida, o cálculo do p-valor foi feito corretamente pelo <i>software</i> , restando a escolha do p-valor uni ou bilateral, o que provocou algum desacordo com o estabelecimento da H_1 , quando implícita.	Considerado um valor a ser comparado com o nível de significância. Não figurou seu caráter probabilístico.
Decisão	Ambos os grupos tiveram um rendimento intermediário, porém, o grupo PBL sim distribuiu-se mais entre os níveis mais baixos e altos do que o grupo PBL não.	Foi comum o encerramento do processo nesta categoria, sem a devida particularização do significado de rejeitar-se ou não a hipótese nula.	Por muitos considerada o ápice do processo, mas nem sempre com o devido entrelaçamento dos conceitos das categorias anteriores, sobressaindo o caráter algorítmico do processo.
Interpretação	O grupo PBL sim teve frequências maiores do que as do PBL não nos extremos dos níveis e o grupo PBL não teve uma concentração maior em níveis intermediários do que o PBL sim.	Categoria com muita diversidade de respostas, inclusive com interpretações que desconsideraram o próprio processo dos THMP, embasando-se na estatística descritiva.	Houve carência na contextualização do processo do THMP com a resposta à situação, entretanto, foram indivíduos do grupo PBL sim que trouxeram as respostas mais adequadas a esta categoria.

Findadas as análises dos dados, passa-se ao capítulo final.

Capítulo 8

Conclusão e Possíveis

Implicações

Esta Tese apresentou, em sua introdução, a necessidade de se olhar o ensino de Estatística para a área da saúde de uma forma mais particular, visto os problemas que esta disciplina vem tendo nos bancos das faculdades e as consequências na pesquisa da área que estas dificuldades podem favorecer.

Na ânsia de prestar um auxílio relevante ao problema, esta pesquisa lançou mão da metodologia ativa do PBL que, amparada pela TCC, compôs as questões de pesquisa normativas deste trabalho e que serão respondidas mais adiante neste texto.

O capítulo 2, sobre a revisão de literatura, pôs em perspectiva o quadro de trabalhos acerca do ensino de Estatística para a graduação em geral e para a área da saúde, em particular. Esta revisão sistemática de literatura levantou achados bem sucedidos e estabeleceu algumas questões carentes de atenção.

Na sequência, pôs-se a fundamentação teórica desta pesquisa e na seção 3.5, apresentou-se a sinergia existente entre os preceitos dados por [Garfield e Ben-Zvi \(2007\)](#), apresentados na revisão de literatura e os ingredientes do PBL, resumidos por [Hadgraft e Holecek \(1995\)](#), além do endosso que a TCC atribui a esta sinergia.

Estando estruturadas as bases deste estudo, houve espaço para o estabelecimento da metodologia de pesquisa (capítulo 4), que encontrou na triangulação metodológica, um formato que permitiu a abordagem de questões importantes do fazer pedagógico, como as impressões do professor/pesquisador (diário de classe, subseção 4.2), o rendimento dos alunos (instrumento fechado, subseção 4.2.2) e o nível de apropriação do CC dos THMP de cada um (instrumento aberto, subseção 4.2.3). Considera-se que a metodologia de pesquisa foi adequada aos propósitos deste estudo.

Na sequência, o capítulo 5, das atividades prévias, permitiu que a metodologia de pesquisa fosse testada e ajustada, além disso, que os processos didáticos fossem melhorados e os conteúdos prévios, trabalhados. Este desenvolvimento autorizou que se partisse para a aplicação definitiva das atividades, o que está descrito no capítulo 6. Das aplicações, depreende-se a necessidade de estruturação do conteúdo dos THMP, bem como do processo do PBL, visto que o alcance natural dos alunos a um conteúdo bastante complexo não se daria somente pela manipulação das situações, senão em um espaço de tempo demasiadamente grande.

Isto feito, atenta-se a lembrar e responder às questões de pesquisa:

- Em que medida se estabelece o campo conceitual dos testes de hipóteses para a média populacional, de alunos da área da saúde, quando expostos ao PBL apoiada pela TCC?

O instrumento para aferição do nível de domínio do CC dos THMP, apresentado na subseção 4.2.3, apontou em sua análise quantitativa que 81% dos participantes atingiu um nível baixo (47%) ou moderado (34%) e apenas 19% atingiu um nível amplo.

A análise qualitativa deste instrumento mostrou que o processo dos THMP pela abordagem do p-valor foi bem compreendido desde o estabelecimento das hipóteses nula e alternativa até a tomada de decisão, passando pela definição do nível de

significância e cálculo do p-valor. Contudo, há carências quanto a interpretação dos resultados e consequente contextualização.

Cabe lembrar que 51 (83,6%) dos 61 alunos participantes do PBL eram do curso de Psicologia, que, como se viu na caracterização da amostra (seção 4.1), mostrou ser um curso avesso ao uso de processos quantitativos, suscitando fragilidades também no uso de softwares como o Excel e grande ansiedade estatística, o que de fato ocorreu, como se viu na descrição do diário de classe (seção 7.1).

- Existe diferença significativa entre o rendimento médio de alunos que experienciaram o PBL apoiada pela TCC e os que não a experimentaram?

O instrumento para aferição de rendimento, mostrado na subseção 4.2.2, apresentou uma diferença de aproximadamente um ponto (em dez pontos possíveis) em favor dos alunos que não participaram do PBL: os que participaram obtiveram uma média de 6,21 pontos com desvio padrão de 1,76 ponto, enquanto os que não participaram tiveram uma média de 7,11 pontos e desvio padrão de 2,11 pontos. Estes valores não se mostraram significativamente distintos ao nível de 5% de significância.

Dado que não houve significância estatística, não há subsídio suficiente para garantir que os alunos não participantes do PBL tenham um rendimento melhor do que os que participaram.

A hipótese de pesquisa foi a seguinte: *A aplicação do PBL apoiado pela TCC cria um ambiente investigativo acerca de situações da área da saúde, nas aulas de bioestatística, provocando a desconforto dos invariantes operatórios dos alunos e sua readequação no decorrer do processo, proporcionando um domínio amplo do campo conceitual dos testes de hipóteses para média de uma amostra, a tal ponto que o rendimento obtido pelos alunos seja compatível ou maior do que daqueles que não passaram pela dinâmica do PBL apoiado pela TCC.*

A parte da hipótese de pesquisa onde se afirma que o rendimento dos alunos participantes do PBL seria compatível ou maior do que os estudantes que tiveram aulas tradicionais não pôde ser validada, não com o endosso da amostra em questão, pois não se percebeu esta diferença a maior, ao contrário, o rendimento dos estudantes participantes do PBL foi menor, embora não significativamente menor, à um nível de significância de 5%, do que o dos não participantes.

A parte que trata do grau de domínio do CC dos THMP dos que participaram do PBL também não pôde ser totalmente validada, a menos para 19% da amostra, o que pode ser considerada uma boa marca, dado o cenário de aplicação da metodologia de ensino, o perfil da amostra e a dificuldade inerente do conteúdo dos THMP.

O ambiente investigativo sobre situações da área da saúde realmente provocou a adequação de esquemas de pensamento bem como a formação de novos esquemas de pensamento (mediante o contato com o conteúdo novo), como visto nos resultados provenientes do diário de classe, o que permite validar esta parte da hipótese de Pesquisa.

Não se pode deixar de reforçar o fato da amostra ser predominantemente proveniente do curso de Psicologia, que não demonstra apreço pelos estudos quantitativos, tal como se auferiu ao caracterizar-se a amostra da pesquisa, na seção 4.1. Outrossim, os demais cursos, embora com tamanhos amostrais bem menores, mostraram desempenhos melhores.

Outras análises são passíveis de serem feitas a partir da experiência vivenciada. Estas, serão dispostas a seguir, considerando algumas categorias que se pôde identificar. Ao final destas análises, vigorarão alguns preceitos que esta pesquisa permitiu auferir, para o melhoramento do ensino de Estatística para a área da saúde.

O PAPEL DO PROFESSOR Desacomodação: se há algo que a metodologia ativa do PBL suscita no professor é a salutar desacomodação, que leva o Professora a repensar, agir diferente, adequar-se, por-se no lugar do aluno, inverter o sentido usual do pensamento do professor. Não é fácil entregar o protagonismo da sala

de aula ao estudante.

Mesmo quando o Professor busca estas mudanças, suas concepções também são resistentes. Se os esquemas de pensamento são persistentes nos alunos, como se depreende de Vergnaud (1990), também o são no professor, mas uma vez que se experimenta tal mudança salutar, não se retrocede ao ponto inicial.

Este efeito aparece também nos colegas Professores da área específica, no caso a área da saúde, que resistiram a auxiliar o professor de Bioestatística durante a aplicação das atividades do PBL, supõe-se que por falta de vontade, desconfiança ou por não entender a metodologia. Esta dificuldade era, de certa forma, previsível, dado que a revisão de literatura trouxe que os professores de estatística em muito provinham da área da matemática ou da estatística, dificultando a integração destes professores aos cursos da área da saúde.

Uma das dificuldades do papel do professor no PBL é a definição da situação principal por, entre outros motivos, ser difícil saber se o problema está adequado ao nível dos alunos. Neste sentido, o conhecimento da zona de desenvolvimento proximal pode ser estabelecido através da pesquisa de invariantes operatórios dos alunos. Esta tarefa difícil, que exige mais de uma técnica de averiguação, pode ser útil se for desenvolvida alguma atividade prévia ao PBL, para que o professor vislumbre sinais destes invariantes antes da definição do problema e do estabelecimento do nível de aprofundamento do conteúdo que se queira trabalhar com a turma.

Ficou claro, dada a experiência vivenciada, que existem pelo menos três momentos destacados do papel do professor: a escolha da situação principal para dar início às atividades, a tutoria da turma durante as atividades e formalização do conteúdo ao final das atividades.

O PBL é bastante dependente da escolha da situação que dá início ao processo

(é o primeiro passo a se estabelecer, segundo Boud e Feletti (1997)) e o professor de Bioestatística precisa contar com a ajuda dos colegas da área da saúde para escolher a situação.

Já a tutoria, depende de um posicionamento mais sensível do professor perante os alunos, na medida em que precisa perceber as necessidades dos alunos para poder orientar-lhes. Neste sentido, a Oportunidade Didática (OD) (definida na subseção 5.1.1) é uma ferramenta de auxílio do professor. Uma das possibilidades para a mudança de um esquema de pensamento equivocado é pela surpresa. A surpresa é um benefício que pode ser aproveitado para a qualificação da OD, principalmente na expansão da Zona de Desenvolvimento Proximal (ZDP) por ruptura de invariante operatório incorreto.

Na OD o professor identifica indícios de invariantes operatórios e a partir deles faz sua intervenção didática para a expansão do conhecimento do aluno, ou seja, não é uma ferramenta didática simples de se manejar e não se sustenta somente pela sensibilidade do professor visto que o conhecimento acerca da TCC é primordial.

A PBL carece de uma finalização do seu processo, não só no sentido de que um estudo não pode durar para sempre, mas também no que concerne à formalização e generalização dos conceitos estatísticos envolvidos no processo e isto só o professor pode fazer. Não é o papel do aluno.

A CONDUTA DO ALUNO A aceitação do PBL por parte dos estudantes nunca foi unânime. Por vezes observou-se muito mais uma resiliência do que motivação. A necessidade de ser diligente quanto à busca dos conceitos nem sempre foi bem vista. A Metodologia ativa não agradou aos alunos tradicionais. A motivação é um objetivo do PBL, mas o aluno tradicional tende a ser impermeável a este objetivo.

A metodologia do PBL tal com adotada nesta pesquisa, teve a vantagem de

receber maior estruturação, o que permitiu acreditar que, aos poucos, a ansiedade do aluno tradicional diminuiu, sem perder o protagonismo do estudante que o PBL exige. O aluno tradicional tem dificuldade de ser protagonista do seu aprendizado. Teve-se a sensação de que a metodologia híbrida proposta foi muito salutar.

O aluno, naturalmente, não quer construir perguntas nem ensinar os colegas. Não transforma os problemas que surgem na pesquisa em questões norteadoras, mas delibera sobre os problemas somente de maneira informal, memorizada, apequenada.

Não raras vezes, os alunos mostraram insegurança no que estão fazendo. Eles têm a ideia de estarem andando em círculos, sem avançarem. Daí a importância do grupo e do tutor.

O aluno, em geral, é tarefairo, precisa de uma ordem sobre o que fazer e dizer se está certo ou não. Situações abertas, como foram, costumam ser trágicas pra eles. O tarefairo calcula, não analisa, não responde e diz que está pronto. Há sempre o risco do grupo dar-se por satisfeito cedo demais com o conhecimento adquirido. Não é fácil para o tutor ultrapassar este muro.

É importante ponderar que a sensação salutar de construção do próprio conhecimento precisa estar pautada pela noção de que a ciência tem um caminho percorrido maior do que o do aluno e que as descobertas dele não são evidências científicas, e raramente serão novas descobertas.

O CONTEÚDO ESPECÍFICO A importância e a necessidade da Estatística na área da saúde é crescente mas o tempo dedicado à Estatística na graduação é pequeno, auxiliando nos problemas de interpretação que a ciência tem sofrido, por conta da formação de um profissional que segue receitas estatísticas pré-definidas mas não pondera adequadamente sobre o método científico estatístico utilizado. Neste sentido, cabe destacar alguns achados:

- O nível de significância é bem entendido como parâmetro de comparação para o p-valor. Neste sentido, a escolha do nível de significância na questão 2 não suscitou o surgimento de novas questões norteadoras (desejáveis durante o PBL), provavelmente por ser um conceito difícil que está atrelado a outros conceitos também difíceis e que, na prática do mecanismo da THMP, serve apenas para a comparação com o p-valor, na visão do aluno.
- Testes de hipóteses são complexos demais para esperar que os alunos os intuem por meio das situações.
- Os alunos tendem a ficar no pensamento amostral e tem dificuldades em pensar inferencialmente.
- Não se pode exigir que os alunos pensem probabilisticamente se, no final das contas, se rejeita ou não, categoricamente, a variável nula. É preciso criar uma nomenclatura que faça jus ao caráter probabilístico dos THMP e que passe essa ideia para os alunos em todos os seus processos.
- Conceitos como o de erro do tipo 1 e 2 e poder do teste, têm a sua importância minorizada em benefício de conceitos mais ligados à tomada de decisão, como o p-valor e o nível de significância.

A ESCOLHA DA SITUAÇÃO PRINCIPAL A escolha do problema gerador da PBL será tanto mais eficiente se permitir desdobramentos em diversas situações nas quais o estudante precisará pôr à prova seus invariantes operatórios para que num processo de acomodação e adequação dos seus esquemas, estabeleçam estruturas compatíveis com o campo conceitual cientificamente aceito.

Problemas escolhidos dentro do contexto do aluno tendem a promover a motivação, auxiliando no processo do PBL, estimulando a explicitação dos seus invariantes operatórios, entretanto, a motivação inicial costuma não se sustentar até o final da atividade sendo sobrepujada pela austeridade da nota.

É importante levar em conta que uma situação aberta pode levar a uma infinidade de outras situações que escapam ao CC que se quer dominar, gerando uma espécie de super desestruturação, passível de promover alguma ansiedade. Mesmo num ambiente melhor estruturado, percebeu-se alguma ansiedade provocada pelo excesso de possibilidades a serem pesquisadas. De outra parte, também corre-se o risco de que a situação não seja encarada como um problema a ser pesquisado, isto ocorreu na aplicação definitiva, onde alguns alunos não viram um questionamento diferente na situação da ESV, que já não haviam estudado na aplicação prévia.

Neste sentido, há sempre o risco de que a situação não suscite a pesquisa que se quer desenvolver pois, como no caso dos THMP, a abordagem à situação é bastante especial e foi desenvolvida ao longo de muitos anos pelos cientistas, sendo pouco provável que uma situação encaminhe, por si só, a esta abordagem especial.

O CONTEXTO DE ENSINO A importância de se proporcionar um ambiente investigativo, de pesquisa, durante o ensino de estatística ficou evidente na revisão bibliográfica (capítulo 2). Situações típicas do processo investigativo, cedo ou tarde, exigirão a desacomodação dos esquemas de pensamento dos estudantes, frente aos novos desafios. O PBL é profícuo neste sentido, entretanto, há que se levar em conta a existência de um campo conceitual subjacente às situações, portanto, o processo investigativo não é de todo aberto, mas obedece os limites mais evidentes do CC.

O PBL tende a ser mais lento do que o ensino tradicional pois não é o professor/tutor que dita o ritmo do desenvolvimento do conteúdo, são os alunos que o fazem. Além disso, é preciso que os alunos entendam o processo da PBL que, para eles, ocorre como uma metodologia de resolução de problemas e não necessariamente uma metodologia de ensino, assim, é premente que eles adotem seus métodos para que o processo realmente ocorra dentro do PBL. Neste sentido, o

papel do tutor é crucial para manter o foco no trabalho auxiliando na sua fluidez ao indicar bibliografia, materiais de consulta, colocando questões que reorientem algum pensamento equivocado do grupo e estruturando o processo na medida do necessário.

Mesmo um processo aberto e focado na pesquisa quanto o PBL pode trazer resultados semelhantes ao modelo tradicional, que costuma ser bastante operacional. Isto por que os alunos pesquisam mais sobre o como fazer, para se salvaguardarem na hora da prova. Aos poucos, os alunos tendem, legitimamente, a se preocuparem com a avaliação. Neste sentido, ter obtido com o PBL, resultados semelhantes aos tradicionais, pode auxiliar os estudantes a perceberem que metodologias ativas dão conta do rendimento que eles procuram.

A ansiedade estatística, pontuada na revisão bibliográfica (capítulo 2), pareceu ser mais difícil de combater do que o apregoadado. Embora não seja o foco desta pesquisa, foi notório que o ambiente investigativo baseado em situações do campo de estudo do aluno, próprio do PBL, não foram mais motivadores do que a nota ao final do semestre.

Talvez, o contexto de ensino onde a pesquisa tomou corpo, descrito na seção 5.2, privilegie o pragmatismo demonstrado pelo aluno dada a quantidade de disciplinas cursadas junto com a Bioestatística, e a quantidade de avaliações geradas por elas. Isto em meio a uma carga horária muito pequena para a disciplina de Bioestatística.

De outra parte, um processo ativo como o do PBL, inserido num contexto tradicional, em um dado momento da disciplina de Bioestatística, irá carecer de tempo para o seu desenvolvimento. Neste sentido, o currículo com três períodos semanais (totalizando 2h e 30min por semana) já é insuficiente para o próprio modelo tradicional, mais rápido.

ÚLTIMAS ASSERTIVAS Algumas análises que fogem das categorias anteriores:

- A PBL diminui o tempo das explicações expositivas pois a necessidade da explicação veio do aluno, melhorando seu interesse na explicação. As OD são mais frequentes neste ambiente no que no tradicional.
- Os alunos aprendem a parte operacional sozinhos por meio do PBL. A parte interpretativa, que é muito maior, até se tem impressão que conseguiriam, mas demoraria muito tempo.
- A PBL é uma metodologia ativa, de cunho construtivista difundida na área da saúde, mas não na disciplina de Bioestatística, que carece de fundamentação teórica em relação a como o aluno aprende, onde a TCC vem auxiliar por meio desta Pesquisa.
- Segundo [Ribeiro \(2008\)](#), um ponto negativo do PBL é a imprecisão no conhecimento das teorias mais avançadas. Este problema fica contornado pela caracterização do campo conceitual cientificamente aceito que a TCC proporciona, deixando claro ao professor, o que falta que os alunos dominem do conteúdo.
- Dependendo do contexto, o PBL sofre com algumas questões:
 - Nem sempre o aluno percebe o conteúdo envolvido na situação;
 - Dificuldades com a generalização do conteúdo;
 - Prevalência de resultados da equipe sobre os individuais;
 - Demora mais que o modelo tradicional;
 - Aprender de forma autônoma pode gerar incerteza acerca do conhecimento adquirido
 - Alta demanda do tutor.
- A dificuldade de se implementar uma metodologia ativa, mesmo flexível e aberta como o PBL, em momentos especiais de currículos tradicionais,

não inviabiliza que o professor adote uma postura que crie um ambiente investigativo, ativo e autônomo em sua sala de aula. Adequações que levem em conta as vantagens das metodologias ativas, são possíveis e desejáveis.

- A situações adicionais, que podem ser vistas no anexo H, poderão ser analisadas em uma próxima pesquisa que objetive o entendimento do processo de domínio do campo conceitual (CC) dos THMP, contrastando com o objetivo desta Pesquisa, que foi o de estabelecer o nível de domínio de tal CC.

Às vezes, talvez sempre, os esquemas de pensamento se formam de maneira intuitiva, uma autogênese do esquema, de tal forma que ele se cria de maneira natural, nem sempre deliberada. Então, quando o aluno está aprendendo, o seu esquema está se estruturando, grande parte das vezes de forma intuitiva, por mais que ele esteja forçando esta criação, estudando. Acontece, muitas vezes, que o professor fala alguma coisa e o aluno entende algumas partes de forma totalmente errada e ali se formam os invariantes operatórios que vão agir mais tarde, a ponto de chegar e dizer “mas o senhor disse”, o que, na verdade, foi uma conclusão que ele tirou de acordo com algo conceitualmente correto que o professor falou e o aluno involuntariamente distorceu e tomou por verdade.

No PBL, este tipo de situação pode se agravar pois, no aprendizado autônomo, não há uma crítica constante e satisfatória do que se está aprendendo, a não ser pelo grupo e pelo tutor. O grupo talvez nem sempre tenha condições suficientes para barrar o conhecimento errôneo e, pode até mesmo, reforçá-lo. Quanto ao tutor, dada a alta demanda gerada pelo PBL, provavelmente não tenha como atender a todas as questões equivocadas que estejam surgindo. Isto reforça a necessidade do professor encerrar o conteúdo por meio de uma formalização e generalização adequados.

Pelo o que se construiu no referencial teórico, e levando em conta a experiência

acumulada ao final desta pesquisa, é importante estabelecer alguns parâmetros passíveis de melhorarem o ensino de Estatística para a área da saúde, quais sejam:

- Professores de Estatística devem inteirar-se com o campo de pesquisa da área da saúde, principalmente quando provenientes de outras áreas do conhecimento.
- As aulas de Estatística devem refletir a pesquisa da área da saúde.
- Leva tempo para os conceitos estatísticos serem compreendidos, por isto, a carga horária da estatística precisa ser a maior possível.
- O ensino de Estatística para a área da saúde deve dar-se em um ambiente investigativo.
- Desenvolvimentos puramente algébricos devem ser preteridos em favor do fortalecimento de conceitos estatísticos. Significados são mais importantes do que resultados.
- Professores de Estatística precisam perceber que valores numéricos, podem transmitir ao aluno a ideia de exatidão matemática, que contrastam com a ideia de variabilidade e com o pensamento probabilístico da Estatística.
- A Estatística da área da saúde deve ser aprendida com o auxílio de computadores.
- O fazer pedagógico do professor de Estatística para a área da saúde precisa estar embasado por uma teoria de aprendizagem.

Com estes parâmetros, acredita-se que aumente a chance de que o ensino de Estatística diminua os sinais de ansiedade estatística, aumente a noção de que a Estatística é crucial para a área da saúde e que há um distanciamento entre Estatística e Matemática, principalmente no que concerne ao estudo da variabilidade, amostragem, inferência e metodologias de pesquisa. Assim, espera-se que a formação do estudante da área da saúde seja aprimorada como um todo, não só no que concerne à Estatística.

8.1 Prolongamentos da Pesquisa

Adaptar uma metodologia ativa, sem descaracterizá-la, a um contexto tradicional de ensino, em um pequeno espaço de tempo e a isso acrescentar algumas das benesses de uma complexa teoria da aprendizagem, foi um desafio que gerou muitos frutos.

Um fruto importante foi a sensação de que não é preciso alterar todo o currículo da disciplina de estatística, nem ocupar tempo demais, para se introduzir um ambiente de investigação em sala de aula. Os resultados mostraram não só que é possível, mas também que as vantagens de um aluno mais ativo em prol do seu aprimoramento gera rendimentos semelhantes ao ensino tradicional, que lhe é familiar.

Para isto, providências de preparação precisaram ser tomadas, como o estudo piloto e as atividades prévias anteriores às atividades principais, o que, ao final, promoveram o desenvolvimento de um grande conjunto de ações do estudante, cujos efeitos perduraram até o final do semestre, mesmo com o retorno da metodologia tradicional.

Outro fruto motivador foi a noção de que se fez algo passível de compartilhamento com os pares professores, que são resistentes à mudanças pois lhes parece desacomodadoras demais, mas com o que foi feito, a realidade de pesquisa do próprio professor passa a ser passível de utilização em sala de aula, em um ambiente menos expositivo e mais participativo com o seu alunos, sem que haja a necessidade de ferramentas, instrumentos ou recursos alheios ao seu conhecimento. Um jogo, por exemplo, pode ser extremamente eficaz no ensino, porém, não costuma fazer parte do cabedal de conhecimento do professor universitário de estatística para a área da saúde e, tende a lhe parecer complexo de utilizar ou infantil demais ou que lhe vai imputar muito trabalho extra, diferentemente da atividade que propusemos nesta pesquisa.

É claro que metodologias ativas exigem que o professor esteja engajado nesta tarefa e que desenvolva questões subjetivas como empatia, motivação, simpatia, comunicação, atenção ao aluno e aos grupos e não só ao conteúdo, entre outras. Estas características

costumam aparecer nos professores atuantes na educação superior que têm a didática em tão alta conta quanto a pesquisa na área específica que desenvolvem.

A partir das experiências acumuladas durante a realização desta Pesquisa, outros desdobramentos dela ficam prementes, quais sejam:

- A primeira das consequências da pesquisa deve ser a replicação das atividades para confirmação, ou não, dos resultados e aprimoramento dos processos.
- A este ponto, tem-se pré-requisitos teóricos suficientes para iniciar, em conjunto com profissionais da área da informática, a concepção de um instrumento aplicável a partir de navegador de internet, responsivo ao estudante (individualizado) mediante sistemas inteligentes (inteligência artificial), que seleciona as questões de acordo com as respostas anteriores e determina de forma autônoma, ao final do teste, em que nível de domínio do CC dos THMP o aluno está, além de dar a estatística da turma toda, para o professor.
- A adaptação de metodologias ativas, mais adequadas a currículos tradicionais, serão propostas, testadas e avaliadas, a fim de que suas vantagens possam chegar até os alunos dos currículos onde as metodologias ativas não são contempladas.
- Mesmo tendo chamado de Aplicação definitiva, ainda é, e provavelmente sempre será, uma aplicação e não a aplicação, visto que uma dificuldade da ciência didática é a constante mudança de todos os elementos envolvidos. A didática não é uma ciência da estabilidade.

Alguns desdobramentos da experiência de Pesquisa já foram contabilizados:

- Elaboração, como professor conteudista, e participação como professor ministrante, em duas edições de um curso de extensão, na modalidade a distância, da Universidade Franciscana, sendo a primeira edição intitulada “Metodologias Ativas

para o Ensino Híbrido”, e a segunda chamado “Metodologias Ativas”. O curso foi produzido em parceria com a Professora Dra. Valéria Iensen Bortoluzzi (1ª edição) e Professora Ms. Ana Paula Pinheiro (2ª edição).

- Muitas análises estatísticas feitas pelo autor desta Pesquisa foram provenientes do contato com os Professores de outros cursos da área da saúde, importantes para o processo do PBL. Algumas foram publicadas, vide, [Pedron et al. \(2019\)](#), [Calcing et al. \(2019\)](#), [Michelotti et al. \(2019\)](#), [Pereira et al. \(2019\)](#) e [Righi et al. \(2016\)](#).
- Alguns capítulos de livros foram escritos: [Biazus et al. \(2020\)](#), [Fioravanti e Mussoi \(2017\)](#) e [Piccolo et al. \(2020\)](#).

O ensino de estatística para a área da saúde, fortemente abordado nesta tese com o resolutivo desejo de aperfeiçoá-lo, pode e deve capilarizar seus efeitos até a população em geral.

Durante a pandemia causada pelo novo coronavírus, iniciada em 2019, a falta de dados sobre o número de contágios causou uma impressão subestimada na população brasileira sobre a gravidade da epidemia. Novamente, as questões inferenciais são capazes de trazer alguma luz num cenário deste tipo, logo, a estatística da área da saúde, pode contribuir para a saúde ao estender-se para a população, mais do que quanto confinada aos profissionais da área. Desta forma, um desdobramento obrigatório deste trabalho é a popularização da estatística da área da saúde.

Referências Bibliográficas

- Amrhein, V., Greenland, S., e McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, 567(7748):305–307.
- ASA Revision Committee (2016). Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016. Technical report.
- Augusto, A. (2014). Metodologias quantitativas/metodologias qualitativas: mais do que uma questão de preferência. *Forum Sociológico*, (24):73–77.
- Banzatto, D. e Kronka, S. (1992). *Experimentação agrícola*. Jaboticabal.
- Batanero, C., Darío Vera, O., e Díaz, C. (2012). Dificultades de estudiantes de Psicología en la comprensión del contraste de hipótesis. *Números: Revista de Didáctica de las Matemáticas*, 80:91–101.
- Batanero, C., Estepa, A., e Godino, J. D. (1991). Análisis Exploratorio de Datos: Sus Posibilidades en la Enseñanza Secundaria. *Suma*, 9:25–31.
- Behrens, J. T. (1997). Principles and Procedures of Exploratory Data Analysis. *Psychological Methods*, 2(2):131–160.
- Ben-Zvi, D. e Garfield, J. (2004). *The challenge of developing statistical literacy, reasoning and thinking*. Kluwer Academic Publishers.

- Biazus, J., Machado, J., Oliveira, L., Blattes, M., Fioravanti, R., Lopes, S., Nardi, T., e Barbieri, Y. (2020). Qualidade de Vida de Jogadores de Futsal. Em *Avaliação, Atividade Física e Saúde 2*. Atena, Ponta Grossa.
- Bonita, R., Beaglehole, R., e Kjellstrom, T. (2010). *Epidemiologia Básica*. Livraria Santos Editora, Santos, 2ª edition.
- Borgatto, A. F. e de Andrade, D. F. (2012). Análise clássica de testes com diferentes graus de dificuldade. *Estudos em Avaliação Educacional*, 23(52):146.
- Boud, D. e Feletti, G. (1997). *The challenge of problem based learning*. Kogan Page.
- Cabral, A. (1996). *Dicionário técnico de psicologia*. Editora Cultrix.
- Calcing, A., Alvarez, G., Patias, L., Machado, Cristina, A., Pedron, F., cristina Moraes, Fioravanti, R., e Moura, D. (2019). A514 Hepatic Profile and Weight Loss After Gastroplasty - 0 and 6 Months of Monitoring. *Surgery for Obesity and Related Diseases*, 15(10):S212.
- Callegari-Jacques, S. M. (2003). *Bioestatística: princípios e aplicações*. Artmed, Porto Alegre.
- Campbell, D. T. e Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio Books.
- Canhota, C. (2008). Qual a importância do estudo piloto? Em Núcleo de Investigação da APMCG, editor, *Investigação Passo a Passo: Perguntas e Respostas para a Investigação Clínica*, capítulo 10, pgs. 69 – 73. APMCG, Lisboa, 1 edition.
- Casella, G. e Berger, R. L. (2011). Inferência estatística-tradução da 2ª edição norteamericana. *Centage Learning*.
- Chernick, M. R. (2011). *The essentials of biostatistics for physicians, nurses, and clinicians*. John Wiley & Sons.

- Cohen, H. W. (2011). P Values: Use and Misuse in Medical Literature. *American Journal of Hypertension*, 24(1):18–23.
- Costa, G. G. d. O. (2012). *Curso de Estatística Inferencial e Probabilidades: Teoria e Prática*. Atlas, São Paulo.
- Cox, D. R. (2006). *Principles of Statistical Inference*. Cambridge University Press.
- Cronbach, L. J. (1951). Coefficient Alpha And The Internal Structure Of Tests. *PSYCHOMETRIKA*, 16(3):297–334.
- da Cruz, D. d. A. L. M. e Pimenta, C. A. d. M. (2005). Prática baseada em evidências, aplicada ao raciocínio diagnóstico. *Rev. Latino-am Enfermagem*, 3(13):415–422.
- da Silveira, F. (2013). Considerações sobre o índice de discriminação de itens em Testes Educacionais. *Educação e Seleção*, (07):54–58.
- Dancey, C. e Reidy, J. (2018). *Estatística Sem Matemática para Psicologia*. Penso Editora, 7 edition.
- Defavori, C. G. e Sarriés, G. A. (2007). A correlação de métodos DEXA e CDEXA em absorptimetria mineral óssea. *Radiologia Brasileira*, 40(3):183–187.
- Delmas, R. C. (2002). Statistical Literacy, Reasoning, and Learning: A Commentary. *Journal of Statistics Education*, 10(3).
- Denzin, N. K. (1978). *The Research Act: A Theoretical Introduction to Sociological Methods*. McGraw-Hill, New York, second edition.
- Díaz, F. R. e López, F. J. B. (2012). *Bioestatística*. Cengage Learning, São Paulo.
- Duch, B. J., Groh, S. E., e Allen, D. E. (2001). The Power of Problem-Based Learning: A Practical "How To" for Teaching Undergraduate Courses in Any Discipline. pg. 274.

- Duncan, B. B., Schmidt, M. I., Giugliani, E. R. J., Duncan, M. S., e Giugliani, C. (2014). *Medicina Ambulatorial-: Conduas de Atenção Primária Baseadas em Evidências*. Artmed Editora.
- Espindola Artola, A., López Benítez, R., Miranda Carbonell, M., Ruiz Socarrás, J. M., e Díaz Garcia, G. M. (2014). Estrategia didáctica para disminuir el estrés académico hacia el contenido estadístico en los estudiantes de medicina. *Humanidades Médicas*, 14(2):499–521.
- Estatcamp (2019). Estimación não paramétrica de densidades: método do núcleo - Análise de Capacidade.
- Fioravanti, R., Greca, I. M., e Villagra, J. A. M. (2019). Caminhos do ensino de estatística para a área da saúde. *Revista Latinoamericana de Investigación en Matemática Educativa*, 22(1):67–96.
- Fioravanti, R., Greca, I. M., Villagra, J. A. M., e Moreira Junior, F. d. J. (2020). Análise do rendimento de estudantes da área da saúde frente a uma metodologia ativa de ensino para testes de hipóteses. *Ciência e Natura*, No prelo.
- Fioravanti, R. e Mussoi, T. D. (2017). Matemática Aplicada à Nutrição. Em *Nutrição - Curso Prático*, capítulo 6, pgs. 410–419. GEN, Rio de Janeiro, 1 edition.
- Fiorentini, D. e Lorenzato, S. (2009). *Investigação em educação matemática: percursos teóricos e metodológicos*. Autores Associados, Campinas, SP, 3^a edition.
- Garfield, J. (1995). How Students Learn Statistics. *International Statistical Review/Revue Internationale de Statistique*, 63:25–34.
- Garfield, J. e Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, 75(3):372–396.

- Gatti, B. A. (2004). Estudos quantitativos em educação. *Educação e Pesquisa*, 30(1):11–30.
- Greca, I. M. e Moreira, M. A. (2003). Do saber fazer ao saber dizer: uma análise do papel da resolução de problemas na aprendizagem conceitual de Física. *Revista Ensaio*, 05:52–67.
- Hadgraft, R. e Holecek, D. (1995). Towards total quality using problem-based learning. *International Journal of Engineering Education*, 11:8.
- Hannigan, A., Hegarty, A. C., e McGrath, D. (2014). Attitudes towards statistics of graduate entry medical students: the role of prior learning experiences. *BMC medical education*, 14:70.
- Herman, A., Notzer, N., Libman, Z., Braunstein, R., e Steinberg, D. M. (2007). Statistical education for medical students - Concepts are what remain when the details are forgotten. *Statistics in Medicine*, (26):4344–4351.
- Hernandez, J. A. E., dos Santos, G. R., da Silva, J. d. O., Mendes, S. L. L., e Ramos, V. d. C. B. (2015). Evidências de Validade da Escala de Ansiedade em Estatística em Alunos da Psicologia. *Psicologia: Ciência e Profissão*, 35(3):659–675.
- Hogg, R. V., Tanis, E. A., e Zimmerman, D. L. (2013). *Probability and Statistical Inference*. Pearson, 9 edition.
- Jones, R., Higgs, R., de Angelis, C., e Prideaux, D. (2001). Changing face of medical curricula. *The Lancet*, 357(9257):699–703.
- Kettley, S. (2020). UK coronavirus death rate: Coronavirus has killed more than SARS - How dangerous is it? *Daily Express*.
- Kuwaiti, A. A. (2015). Health Science students' evaluation of courses and Instructors:

- the effect of response rate and class size interaction. *International Journal of Health Sciences*, 9(1):51–60.
- Maroco, J. e Garcia-Marques, T. (2006). Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas? *Laboratório de Psicologia*, 4(1):65–90.
- Martinez, E. Z. (2015). *Bioestatística para os cursos de graduação da área da saúde*. Blucher, São Paulo.
- Mayer, D. (2009). *A brief history of medicine and statistics*. Cambridge University Press, 2nd edition.
- Michelotti, T. C., Fioravanti, R., e de Moraes, C. M. B. (2019). Qualidade da Dieta de Universitários. Em *6º Congresso Internacional em Saúde*, number 6.
- Moreira, M. A. (2002). A TEORIA DOS CAMPOS CONCEITUAIS DE VERGNAUD, O ENSINO DE CIÊNCIAS E A PESQUISA NESTA ÁREA. *Investigações em Ensino de Ciências*, 7(1):7–29.
- Moreira, M. A. (2011). *Metodologias de Pesquisa em Ensino*. São Paulo, 1 edition.
- Morettin, P. A. e Bussab, W. O. (2017). *Estatística básica*. Editora Saraiva.
- Nogueira, C. M. I. e Rezende, V. (2014). A Teoria dos Campos Conceituais no Ensino de Números Irracionais : Implicações da Teoria Piagetiana no Ensino de Matemática. *Revista Eletrônica de Psicologia e Epistemologia Gnética*, 6(1):41–63.
- Nuzzo, R. (2014). Statistical Errors: P values, the ‘gold standard’ of statistical validity, are not as reliable as many scientists assume. *Nature*, 506:150–152.
- Oliver, A., Sancho, P., Galiana, L., e Cebrià i Sancho, M. Á. (2014). Nueva evidencia sobre la Statistical Anxiety Scale (SAS). *Anales de Psicología*, 30(1):150–156.
- Orange Data Mining (2019). Distributions.

- Pagano, M. e Gauvreau, K. (2006). *Princípios de Bioestatística*. Thomson Learning, São Paulo.
- Panagiotakos, D. B. (2008). Value of p-value in biomedical research. *The open cardiovascular medicine journal*, 2:97–9.
- Patton, M. Q. (1980). *Qualitative evaluation methods*. SAGE, Beverly Hills, 1 edition.
- Pedron, F., Alvarez, G., Patias, L., Calcing, A., Moura, D., Machado, Cristina, A., Moraes, C., e Fioravanti, R. (2019). A515 Loss of Weight, Glucose and Hemoglobin Glyced in Patients with Diabetes Mellitus II - 6 Months after Gastroplasty. *Surgery for Obesity and Related Diseases*, 15(10):S212–S213.
- Pereira, G. d. S., Mussoi, T. D., e Fioravanti, R. (2019). Estado nutricional e percepção da imagem corporal de acadêmicas universitárias do curso de Nutrição do Brasil e de Portugal. *Disciplinarum Scientia*, 20(1):85–94.
- Piccolo, A. M., Rossato, D. D., Biazus, J. d. F., Machado, J. R. S., Lilian Oliveira de Oliveira, Blattes, M. W., Fioravanti, R., e Gomes, T. J. N. (2020). Qualidade do Sono como Preditor de Lesões Musculares em Jogadores de Futebol Profissional de um Clube de Santa Maria/RS. Em *Ciências da Saúde: Campo Promissor em Pesquisa 4*, capítulo 19, pgs. 183–191. Atena, Ponta Grossa, 1 edition.
- Post, W. J. e Van Duijn, M. A. J. (2014). Teaching Hypothesis Testing: a Necessary Challenge. Em *ICOTS9*, pgs. 1–3, Flagstaff, USA.
- Regnier, J.-C. (2004). DA VERDADE AUTO-PROCLAMADA À VEROSIMILHANÇA RECONHECIDA : UM PONTO CENTRAL NA FORMAÇÃO EM ESTATÍSTICA. Em *VIII Encontro Nacional de Educação Matemática*, pg. 29, Recife.
- Ribeiro, L. R. d. C. (2008). Aprendizagem baseada em problemas (pbl): uma experiência no ensino superior.

- Righi, A. S., Basso, C., e Fioravanti, R. (2016). Rejeitos alimentares em um hospital filantrópico da região Centro-Oeste do estado do Rio Grande do Sul. *Disciplinarum Scientia*, 17(2):267–277.
- Rizopoulos, D. (2018). Latent Trait Models under IRT (ltm). Technical report, CRAN(R-project).
- Savin-Baden, M., Howell, C., Savin-Baden, M., e Major, H. (2004). *Foundations of Problem-based Learning*. McGraw-Hill, Maidenhead-UK.
- Schmidt Godoy, A. (1995). Introdução à Pesquisa Qualitativa e suas Possibilidades. *Revista de Administração de Empresas*, 35(2):57–63.
- Sebastiani, R. G. (2011). Teste de Hipóteses: uma análise dos erros cometidos por alunos de engenharia Hypothesis Test : an analysis of errors committed by. *Bolema*, 24(40):835–854.
- Smith, K. L. (2010). Divergent Needs of Learners in Evidence Based Medicine.
- Soares, J. A. R., Amorim, A. F., e da Silva, C. R. (2018). Avaliação Educacional em Larga Escala e Algumas Considerações Sobre a TCT e a TRI. *Revista Ciências Exatas e Naturais*, 20(1):119–125.
- Teresita Terány e Jesica Ciminari (2019). Los errores en la resolución de situaciones problemáticas de inferencia estadística como principal obstáculo en la didáctica de la estadística. Em *Congreso International Virtual de Educación Estadística.*, Rosario.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley, 1st edition.
- Turik, C., Viali, L., e de Moraes, J. F. D. (2012). Análise de atitudes de alunos universitários em relação à estatística por meio da teoria de resposta ao item. *Ciência & Educação (Bauru)*, 18(1):231–243.

- Urbina, S. (2009). *Fundamentos da testagem psicológica*. Artmed Editora.
- Valente, J. A., de Almeida, M. B. E., e Geraldini, A. F. S. (2017). Metodologias ativas: das concepções às práticas em distintos níveis de ensino. *Revista Diálogo Educacional*, 17(52):455–478.
- Vergnaud, G. (1990). Epistemology and Psychology of Mathematics Education. Em Nesher, P. e Kilpatrick, J., editores, *Mathematics and Cognition*, pgs. 14–30. Cambridge University Press, Cambridge.
- Vergnaud, G. (1998). Towards a Cognitive Theory of Practice. Em Sierpinska, A. e Kilpatrick, J., editores, *Mathematics Education as a Research Domain: A Search for Identity*, pgs. 227–242. Springer Science, New York.
- Vergnaud, G. (2003a). A Gênese dos Campos Conceituais. Em Grossi, E. P., editor, *Por que ainda há quem não aprende? A Teoria*, pgs. 21–60. Editora Vozes, Petrópolis.
- Vergnaud, G. (2003b). A psicologia da educação. Em *As Ciências da educação*, capítulo IV, pgs. 63–79. Edições Loyola, S.
- Vergnaud, G. (2008). *Atividade humana e conceitualização*. Porto Alegre.
- Vergnaud, G. (2009). The theory of conceptual fields. *Human Development*, 52(2):83–94.
- Vergnaud, G. (2017). O que é aprender? Por que Teoria dos Campos Conceituais? Em Esther Pillar Grossi, editor, *O que é aprender? Iceberg da conceitualização*, pgs. 17–35. Geempa, Porto Alegre.
- Vigil-Colet, A., Lorenzo-Seva, U., e Condon, L. (2008). Development and Validation of the Statistics Anxiety Scale. *Psicothema*, 20(1):174–180.
- Wasserstein, R. L. e Lazar, N. A. (2016). The ASA statement on p-values: context, process, and purpose. *The American Statistician*, 70(2):129–133.

Yew, E. H. J. e Goh, K. (2016). Problem-Based Learning: An Overview of its Process and Impact on Learning.

Apêndice A

Plano de Ensino da Bioestatística

Plano de Ensino típico dos cursos da área da saúde (Biomedicina, Farmácia, Fisioterapia, Medicina, Nutrição, Odontologia, Psicologia e Terapia Ocupacional) da instituição onde foram aplicadas as atividades da Pesquisa. A seguir, a título de exemplo, apresentam-se a primeira e última páginas do Plano de Ensino do curso de Farmácia/2018.

Plano de ensino

1) Identificação	
Curso	MTM213 FARMÁCIA
Carga horária	34h
Disciplina	BIOESTATÍSTICA
Semestre letivo	GRADUAÇÃO 2018/2
Professor	RODRIGO FIORAVANTI PEREIRA

2) Objetivos
Promover a aquisição de conhecimentos sobre a descrição, a análise e a interpretação de dados biomédicos. Estudar os principais testes estatísticos e dominar a sua aplicabilidade utilizando softwares especializados.

3) Conteúdo Programático
Papel da estatística na área da saúde. Organização da pesquisa. Medidas de posição e dispersão. Caracterização estatística de variáveis. Amostragem. Estimativa estatística. Testes de hipóteses. Correlação e regressão. Testes não-paramétricos. Determinação do teste estatístico a utilizar.

4) Caracterização geral da metodologia de ensino
As necessidades pedagógicas para o ensino de estatística vêm aumentando na medida em que o pensamento estatístico é cada vez mais necessário às ciências da saúde. Neste sentido, o estudante desta área precisa desenvolver habilidades estatísticas que vão do conteúdo teórico ao resultado prático, passando pelas ferramentas tecnológicas utilizadas para o tratamento adequado dos dados. Para cumprir estes objetivos, a disciplina de Bioestatística MTM 364, é desenvolvida em dois momentos semanais, um chamado de teórico (em sala de aula) e outro prático (no laboratório de informática). Os dois momentos procuram desenvolver o conteúdo de forma simultânea para que fique evidente que não há separação entre teoria e prática. As aulas são baseadas em situações que evidenciam a necessidade da estatística na tomada de decisões, com o amparo de ferramentas computacionais que dão conta do caráter algébrico da disciplina. Desta forma a capacidade de interpretação, comunicação escrita e argumentação são privilegiadas, em detrimento de sua capacidade puramente algébrica.

5) Cronograma de desenvolvimento	
Data	Conteúdo/atividade docente e discente
31 julho	Conteúdo: Apresentação do programa, bibliografia, metodologia e datas das avaliações através da discussão do Plano de Ensino. Introdução ao método e conceitos de análise estatística na área biomédica com exemplos práticos. Apresentação das normas de utilização do laboratório de informática. Apresentação dos programas estatísticos e seus principais comandos. Fonte de referência: Apostila disponibilizada pelo professor. Atividades: Aula expositiva com exercícios resolvidos no quadro. Aula prática com o uso de programa estatístico no laboratório de informática.

A avaliação da aprendizagem permitirá que o aluno demonstre conhecer os conceitos fundamentais da Bioestatística e aplicar os princípios teóricos utilizando programa(s) estatístico(s).

Instrumentos de Avaliação:

Média do Primeiro Bimestre = Trabalhos (10,0)

Média do Segundo Bimestre = Trabalhos (10,0)

Média Semestral = (Média do Primeiro Bimestre + Média do Segundo Bimestre) / 2

A aprovação na disciplina de Bioestatística, somente é alcançada, se o estudante:

- Obtiver nota igual ou superior a sete nas avaliações (média semestral), com a frequência mínima de 75%.

- Obtiver nota igual ou superior a cinco calculados pela média aritmética da nota semestral e da nota da avaliação final (exame), com frequência mínima de 75%.

Observação 1: O aluno que faltar a qualquer avaliação deverá justificar conforme as normas que constam do Guia Acadêmico da PROGRAD.

Observação 2: As datas das provas atrasadas serão marcadas fora do horário normal de aula.

7) Bibliografia básica

CALLEGARI-JACQUES, S. M. Bioestatística: princípios e aplicações. Porto Alegre: Artmed, 2003.

RODRIGUES, P. C. Bioestatística. 3. ed. Niterói: Eduff, 2002.

VIEIRA, S. Bioestatística: tópicos avançados. Rio de Janeiro: Campus, 2003.

_____. Introdução à bioestatística. 3. ed. Rio de Janeiro: Campus, 1980.

8) Bibliografia complementar

LUIZ, R. R.; COSTA, A. J. L.; NADANOVSKY, P. Epidemiologia & Bioestatística em Odontologia. São Paulo: Atheneu, 2008.

MOTTA, V. T.; WAGNER, M. B. Bioestatística. São Paulo: Robe, 2003.

SOARES, J. F.; SIQUEIRA, A. L. Introdução à Estatística Médica. 2. ed. Belo Horizonte:

COOPMED, 2002.

WEYNE, G. R. Bioestatística e validade de trabalhos científicos. 2. ed. São Paulo: Scortecchi, 2009.

Apêndice B

Caracterização dos Cursos de Biomedicina, Nutrição e Psicologia

As caracterizações descritas abaixo foram colhidas do site da instituição de ensino que recebeu esta Pesquisa.

Biomedicina Biomedicina é a área da saúde voltada para a pesquisa das doenças humanas, suas causas e os meios de diagnosticá-las. O profissional é um pesquisador capacitado a atuar em muitos níveis de atenção à saúde e é o grande responsável pela garantia dos exames laboratoriais e das análises clínicas.

O curso de graduação em Biomedicina da Universidade Franciscana aborda em seu currículo o diagnóstico laboratorial de biologia molecular, genética e análises clínicas. Para tanto, possui dois estágios de 510 horas (cada), permitindo ao aluno maior aperfeiçoamento no diagnóstico laboratorial e na área de pesquisa.

Os temas de estudo são provenientes das ciências biológicas da saúde humana, ciências humanas e sociais, ciências exatas e ciências da biomedicina. Essas áreas

são permeadas, continuamente, pelos princípios e diretrizes do Sistema Único de Saúde – SUS.

O biomédico, no exercício das atribuições, poderá assumir responsabilidade técnica de laboratórios, de indústrias e firmar os respectivos laudos e pareceres. Para atuar no mercado de trabalho, o mesmo deverá ter o reconhecimento de habilitação na área específica e prévia inscrição no Conselho Regional de Biomedicina.

Nota do Enade: 4

Conceito do MEC: 4

Duração: 8 Semestres

Turno: Noite

Situação Legal: Portaria nº 134, de 1/03/2018 | DOU de 02/03/2018 | Edição 42 | Seção 1 | Página 74.

Início de Funcionamento: 08/11/2006

Habilitação: Bacharel

Nutrição O nutricionista é responsável por investigar e controlar a relação do homem com seu alimento, para assim preservar a saúde humana. As áreas de estudo que fundamentam a profissão são provenientes das ciências biológicas e da saúde; ciências sociais, humanas e econômicas; ciências da alimentação, nutrição e ciências dos alimentos.

O curso de graduação em Nutrição da Universidade Franciscana tem inserido no mercado de trabalho em serviços públicos e privados em todo país. O curso é comprometido em formar profissionais capacitados na promoção, proteção, recuperação da saúde e prevenção de doenças de indivíduos ou grupos populacionais, em todas as etapas da vida.

Nota do Enade: 3

Conceito do MEC: 3

Duração: 8 Semestres

Turno: Tarde

Situação Legal: RENOVADO O RECONHECIMENTO DO CURSO DE NUTRIÇÃO, BACHARELADO, PELA PORTARIA Nº 134, DE 01 DE MARÇO DE 2018, PUBLICADO NO D.O.U DE 02 DE MARÇO DE 2018.

Início de Funcionamento: 03/08/1998

Habilitação: Bacharel

Psicologia O curso de graduação em Psicologia da Universidade Franciscana está organizado em quatro ênfases de competências e habilidades, definidas pelas atuais demandas sociais e condições da instituição. São elas: psicologia comunitária e da saúde; psicologia e processos organizacionais e do trabalho; psicologia e processos clínicos; psicologia e processos educacionais.

No curso, a responsabilidade é a formação de profissionais comprometidos com a prevenção e promoção da saúde nos diferentes contextos de atuação profissional. As disciplinas optativas também são ofertadas de acordo com o interesse e a necessidade dos estudantes. Além disso, a formação do estudante não se resume ao espaço de sala de aula, pois o quadro de professores é diversificado e oferece oportunidades no campo da pesquisa, da extensão e em projetos culturais.

Em adição, os acadêmicos realizam estágios no âmbito jurídico, em escolas, empresas, na clínica, nas comunidades, em hospitais, instituições de acolhimento, ONGs, que permitem ampla experiência em quatro diferentes ênfases: psicologia e processos de prevenção e promoção da saúde; processos clínicos; processos organizacionais e do trabalho e processos educacionais.

Dessa forma, os conhecimentos, habilidades e competências para a formação desse

perfil de psicólogo têm o potencial de garantir a oportunidade de ingresso em um campo de atuação bastante vasto, bem como de atuar em áreas mais específicas.

Nota do Enade: 3

Conceito do MEC: 4

Apêndice C

Mapa Conceitual do CC dos THMP

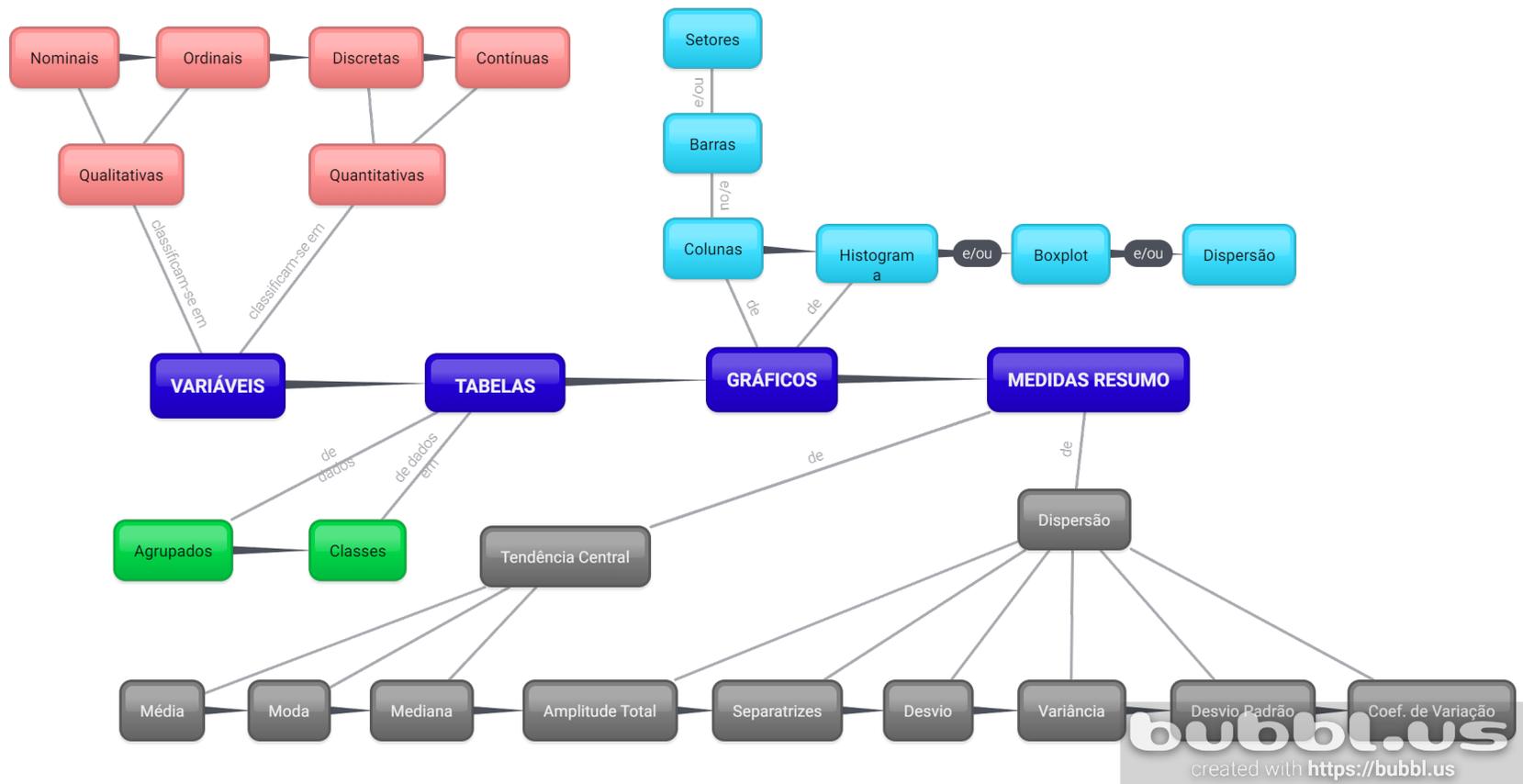
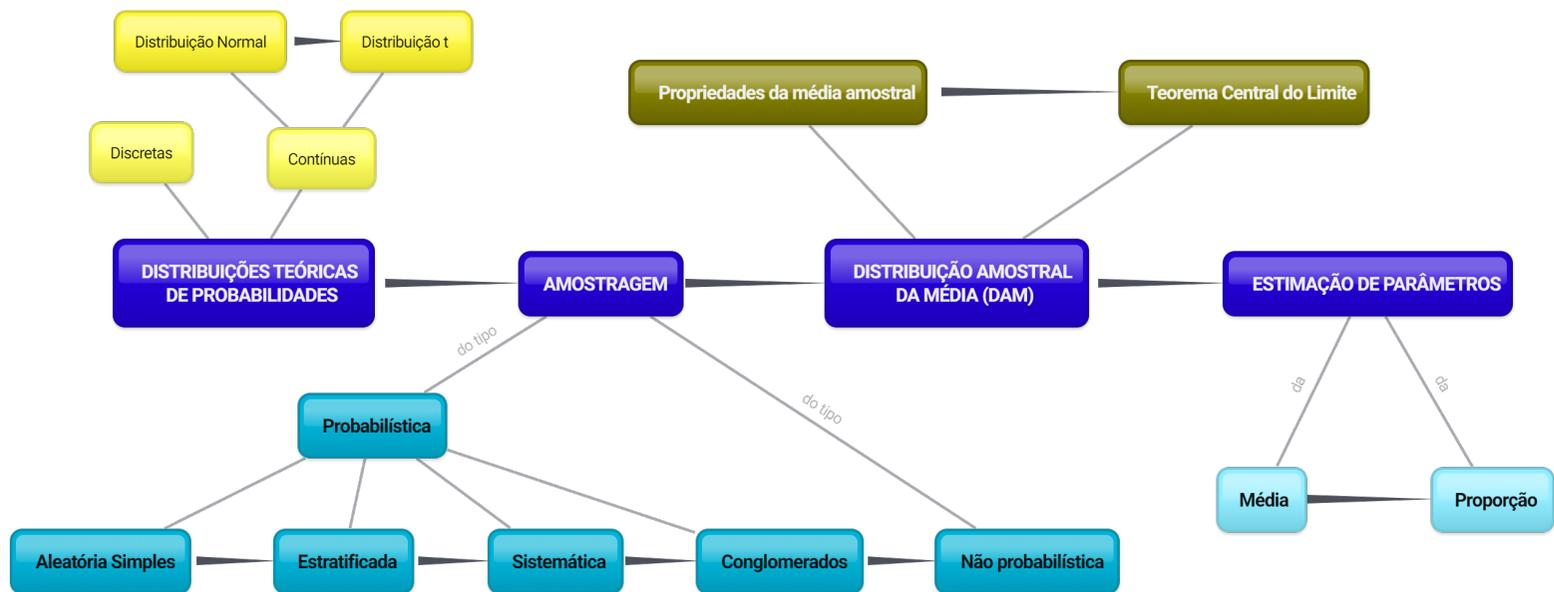
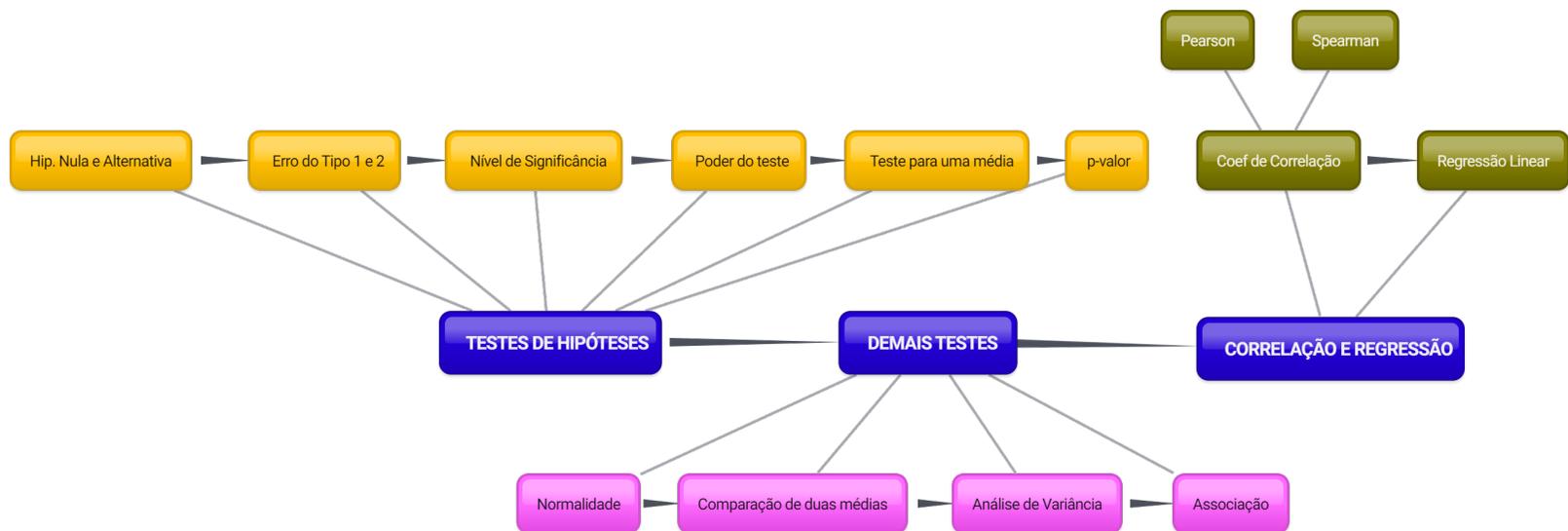


Figura C.1: Mapa conceitual do campo conceitual de Teste de Hipóteses para a Média Populacional (THMP).



created with www.bubbl.us

Figura C.2: Mapa conceitual do campo conceitual de Teste de Hipóteses para a Média Populacional (THMP). (Cont.)



created with www.bubbl.us

Figura C.3: Mapa conceitual do campo conceitual de Teste de Hipóteses para a Média Populacional (THMP). (Cont.)

Apêndice D

Capturas de Telas do Ambiente

Moodle

TESTE DE HIPÓTESES

Através da PBL (Problem-Based Learning)

O uso dos intervalos de confiança representa uma forma bastante informativa de apresentar os achados principais de um estudo que faz uso de amostras, como no caso da estimação da média, em que esses resultados são extrapolados para a população que originou os dados. Em alguns estudos, pode haver o objetivo de *testar hipóteses formuladas sobre parâmetros de interesse*. O teste de hipótese é um procedimento de inferência estatística usual nesses casos. Intervalos de confiança e testes de hipóteses estão intimamente relacionados, dado que seus fundamentos teóricos, em grande parte, são os mesmos. Entretanto, os testes de hipóteses têm algumas características importantes quanto à interpretação de seus resultados.

Uma hipótese estatística é uma suposição sobre um parâmetro de interesse e um teste estatístico de hipóteses é uma regra utilizada para decidir quando rejeitar uma hipótese, sendo essa regra sempre baseada em uma amostra aleatória.

Situação Problema

O arquivo "Diabete mulheres 21+" contém uma série de variáveis com dados extraídos de uma amostra aleatória de mulheres maiores de 21 anos.

Os pesquisadores se depararam com importantes questões acerca da população das mulheres com mais de 21 anos:

Considerando que o nível de glicose dito normal deve estar abaixo de 99 mg/dl, a amostra que temos, permite aceitar a hipótese de que as mulheres desta população tem uma glicose normal?

Outras perguntas surgem desta, por exemplo:

- Qual a hipótese pode ser formulada nesta situação?
- A amostra permite aceitar ou rejeitar esta hipótese? Sob quais condições se aceita ou rejeita a hipótese?

 [Diabete mulheres 21+](#)

Figura D.1: Captura de tela do Ambiente Moodle para atividades do estudo piloto.

-  [O que são Testes de Hipótese?](#)
-  [O que são Testes de Hipóteses | Para que servem os Testes de Hipóteses](#)
-  [As Hipóteses Nula e Alternativa](#)

Este material é importante para complementar a aula anterior e responder a algumas questões levantadas, além de preparar a próxima aula.

Material extraído do livro Bioestatística para os Cursos de Graduação da Área da Saúde, de Edson Zangiacomi Martinez.

Figura D.2: Materiais de apoio oferecido aos estudantes durante o estudo piloto.

 Esquema básico dos Testes de Hipóteses

Assista ao vídeo e prepare-se para a próxima aula!

A imagem abaixo é o roteiro que aparece no vídeo.

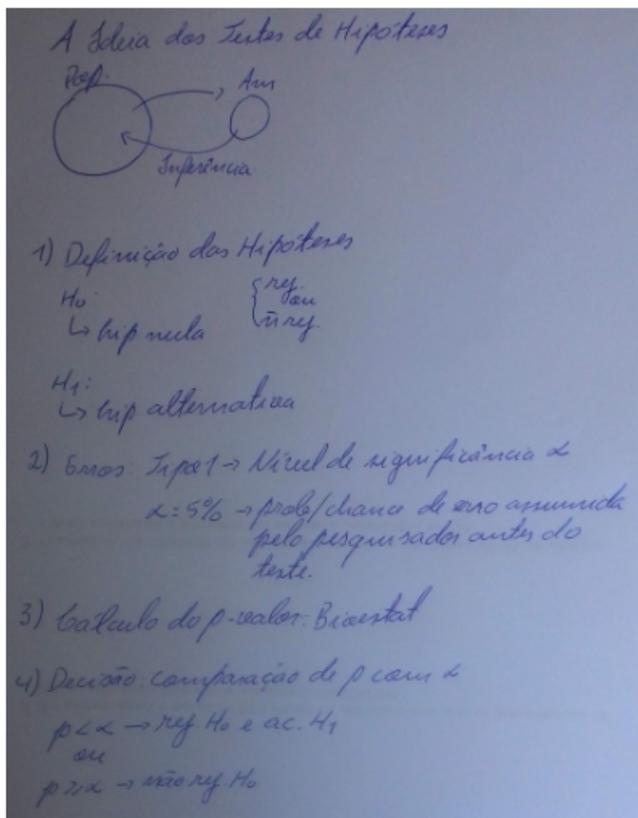


Figura D.3: Cabeçalho e primeiros dados da base de dados do estudo piloto.

Apêndice E

A Escala de Satisfação de Vida (ESV)

CAPÍTULO DE LIVRO COM A ESCALA DE SATISFAÇÃO DE VIDA

Fonte: Avaliação em Psicologia Positiva. Claudio Simon Hutz (Org.). Artmed, 2014.

Satisfação de vida é o componente cognitivo do bem-estar subjetivo definido como o nível de contentamento que alguém percebe quando pensa sobre sua vida de modo geral. Em outras palavras, a satisfação de vida pode ser entendida como o nível de entusiasmo e prazer, ou descontentamento e sofrimento, presente na vida de uma pessoa de acordo com sua percepção do que é satisfatório e/ou desprazeroso (Diener, Lucas & Oishi, 2005). Por isso, diz-se que é uma avaliação subjetiva realizada em momentos em que há uma autorreflexão sobre aspectos importantes da vida. Essa avaliação leva em conta amplos aspectos da vida como um todo e tende a ser estável ao longo do tempo (Lyuburnirsky, King & Dierier, 2005).

Para avaliar sua satisfação de vida, as pessoas considerarão tanto coisas boas quanto coisas ruins. Alguns priorizarão eventos agradáveis, enquanto outros focarão a atenção em eventos desagradáveis. Esse julgamento sobre o quão satisfeito se é com a vida

reflete informações diferentes para pessoas distintas e pode mudar dependendo de humor, eventos de vida, pensamentos e sentimentos presentes no momento (Dierier, Lucas & Oishi, 2005). Contudo, aquilo que consideramos importante tende a não mudar, e há evidências de que as lembranças mais salientes na memória apresentam considerável estabilidade na hora de avaliar a satisfação de vida (Kuppens, Realo & Diener, 2008).

A satisfação de vida pode ser explicada em grande parte pela genética e pelos traços de personalidade (Diener & Lucas, 2010). Estudos com gêmeos monozigóticos (idêntica carga genética) e heterozigóticos (carga genética diferente) revelaram que os níveis de bem-estar eram semelhantes para os monozigóticos, mas diferentes para os heterozigóticos. Esses resultados se mantiveram mesmo controlando participantes que foram criados por famílias diferentes (Lykken & Tellegen, 1996; Nes, Roysamb, Tambs, Hanis & Reichbornenrend, 2006). De fato, a carga genética parece determinar em grande parte a satisfação de vida de gêmeas monozigóticos criados separadamente. Isso corrobora a hipótese de que o meio ambiente tem pouca influência no bem-estar, enquanto a herança genética apresenta grande impacto sobre ele.

A personalidade também parece ser um importante preditor de bem-estar (e consequentemente de satisfação de vida). Especialmente os fatores neuroticismo e extroversão são muito relevantes para o modo como as pessoas experienciam a felicidade em suas vidas (Steel, Schmidt, & Shultz, 2008). De forma geral, pessoas mais ansiosas, inseguras, depressivas, instáveis emocionalmente, ou seja, com altos níveis de neuroticismo, tendem a vivenciar muitas emoções, sentimentos, pensamentos negativos e ver o mundo de forma mais opressiva. Já pessoas mais comunicativas, com muitos contatos sociais, assertivas, enérgicas, ou seja, com altos níveis de extroversão, tendem a sentir mais emoções e sentimentos positivos, além de apresentar uma visão mais positiva sobre o mundo. Por isso, aquelas com altos níveis de neuroticismo sentem-se mais insatisfeitas com suas vidas, enquanto as com altos níveis de extroversão sentem-se mais satisfeitas (Lucas & Diener, 2010). Apesar da grande importância da genética e da personalidade para

a satisfação de vida, outras variáveis, como facilidade para recordar eventos positivos ou negativos, esperança e otimismo, também são relevantes (Diener, Lucas, & Oishi, 2005). Em outras palavras, diferenças individuais para acessar informações, assim como disposições cognitivas diversas, estão associadas à satisfação de vida. Ou seja, evocar mais, e mais facilmente, informações positivas, apresentar expectativas positivas sobre o futuro (otimismo), assim como buscar ativamente aquilo que se considera importante (esperança) também parece influenciar os níveis de satisfação de vida (Diener, Lucas, & Oishi, 2005).

Uma escala mundialmente usada para medir o componente cognitivo do bem-estar subjetivo é a Escala de Satisfação de Vida (ESV; Diener, Emmons, Larsen, & Griffin, 1985). Essa escala foi adaptada e validada para adultos e adolescentes brasileiros pelo grupo de pesquisadores do Laboratório de Mensuração da UFRGS (ver, p. Zanon et al., 2013) e tem sido usada em inúmeras pesquisas no País, inclusive em dissertações de mestrado e teses de doutorado.

NORMAS PARA APLICAÇÃO DA ESV

A ESV é composta de cinco itens de autorrelato, cujo conteúdo avalia o nível de satisfação dos sujeitos com suas condições de vida. A chave de respostas é uma escala Likert de sete pontos em que as pessoas assinalam um número que corresponde ao quanto concordam ou discordam das sentenças apresentadas. As ancoras "1" e "7" recebem os valores "Discordo plenamente" e "Concordo plenamente", respectivamente, enquanto os demais valores intermediários representam diferentes níveis de concordância/discordância com os itens. Quanto mais próximo de "1", mais o sujeito discorda, e, quanto mais próximo de "7", mais ele concorda com a sentença. Instruções são dadas ao respondente no cabeçalho da escala.

ESCALA DE SATISFAÇÃO DE VIDA

Laboratório de Mensuração da UFRGS

Instruções

Abaixo você encontrará cinco afirmativas. Assinale na escala abaixo de cada afirmativa o quanto ela descreve a sua situação pessoal. Não há respostas certas ou erradas, mas é importante você marcar com sinceridade como você se sente com relação a cada uma dessas afirmativas.

1) A minha vida está próxima do meu ideal.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

2) Minhas condições de vida são excelentes.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

3) Eu estou satisfeito com a minha vida.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

4) Até agora eu tenho conseguido as coisas importantes que eu quero na vida.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

5) Se eu pudesse viver a minha vida de novo eu não mudaria quase nada.

Discordo plenamente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Concordo plenamente

Figura E.1: Escala de Satisfação de Vida.

NORMAS PARA LEVANTAMENTO E INTERPRETAÇÃO DA ESV

Para conhecer o nível de satisfação de vida da pessoa que respondeu à escala, inicialmente devem-se somar todos os itens respondidos. O valor obtido representa o escore bruto. Posteriormente, procura-se na tabela de normas adequada qual é o percentil correspondente ao escore bruto. Quanto mais alto o percentil equivalente ao escore bruto, maior a satisfação de vida. Por exemplo: um adulto que obteve escore bruto igual a 9 está no percentil 5 (Tab. 1), está junto com os 5% de sujeitos que apresentaram os menores escores da escala, ou menor satisfação de vida; enquanto uma pessoa que apresentou escore bruto de 22 está no percentil 45 (o que significa que apresentou níveis médios de satisfação com a vida). Já um adolescente que apresentou escore bruto de 33 está no percentil 95 (Tab. 2), o que significa dizer que seu escore

está junto aos 5% dos adolescentes mais satisfeitos. Esses valores refletem o nível de satisfação do sujeito quando comparados à amostra de normatização da escala. A seguir são fornecidas normas para interpretação dos escores de adultos (Tab. 1) e adolescentes (Tab. 2).

Tabela 1
NORMAS DA ESCALA DE SATISFAÇÃO DE VIDA PARA HOMENS E
MULHERES

Percentil	Escore brutos	Escore T
5	9	32
10	11	35
15	13	38
20	15	41
25	17	43
30	18	45
35	19	46
40	21	49
45	22	50
50	23	52
55	24	53
60	25	54
65	26	56
70	27	57
75	28	58
80	29	60
85	30	61
90	31	62
95	32	64
	M 21,8	
	DP 7,3	

Tabela 2

NORMAS DA ESCALA DE SATISFAÇÃO DE VIDA PARA ADOLESCENTES DE
AMBOS OS SEXOS

Percentil	Escore brutos	Escore T
5	12	32
10	14	35
15	16	38
20	18	42
25	19	43
30	20	45
35	21	46
40	22	48
45	23	50
50	24	51
55	25	53
60	26	55
65	26	56
70	27	57
75	28	58
80	29	59
85	30	61
90	31	63
95	33	66
	M 23,2	
	DP 6,2	

Apêndice F

Situação e Material de Apoio para Aplicação Definitiva: Curso de Psicologia

TESTES DE HIPÓTESE PARA A MÉDIA POPULACIONAL (THMP)



Nesta segunda incursão no PBL, os grupos vão analisar a seguinte situação:

O Professor de Estatística quer avaliar o nível de Satisfação de Vida dos alunos de Estatística de sua Universidade. Para isto aplicou a Escala de Satisfação de Vida (ESV) para os alunos que se voluntariaram, o que gerou a amostra cujos dados estão no arquivo "ESV todos os alunos com escore".

As perguntas a seguir ajudarão no processo de entendimento de um Teste de Hipótese para uma Média:

1. Qual é a média da amostra?
2. Considerando que um escore 25 na ESV indica que se está satisfeito com a vida. Existe alguma chance (probabilidade) da população ter média 25, mesmo que a amostra não tenha média 25? Esta chance é grande ou pequena?
3. Se a população tiver média 25, existe alguma chance de se retirar uma amostra cuja média não é 25?
4. Se a população tiver média 25, qual a chance de tirar uma amostra com a média que obtivemos? Esta chance é grande ou pequena? O que significa grande ou pequeno neste contexto?
5. Conduza um teste de hipóteses para uma média com base na amostra levantada e decida se a média populacional é de 25.
 - a) Estabeleça a hipótese a ser testada e a alternativa. Escreva, com suas palavras, o que elas significam dentro do contexto desta situação.
 - b) Defina o nível de significância, justifique a sua escolha e explique o seu significado dentro do contexto da questão.
 - c) Defina a região de aceitação e rejeição de H_0 .
 - d) Calcule a estatística do teste e a localize no gráfico desenhado em c).
 - e) Calcule o valor de p adequado e explique o seu significado dentro do contexto da questão.
 - f) Decida sobre rejeitar ou não a hipótese nula e estruture uma resposta escrita para a situação.

Não esqueça que na PBL, o grupo analisa a situação, faz um levantamento do que já sabe e do que é necessário aprender (questões de aprendizagem), distribui as questões de aprendizagem para que os integrantes possam pesquisá-las e compartilharem o que descobriram no próximo encontro, procurando solucionar a situação.

 [ESV todos os alunos com escore](#)

MATERIAL DE APOIO

O material a seguir ajudará o grupo a conduzir os Testes de Hipóteses apropriados para resolver as questões da situação.

A sequência de vídeos abaixo dá uma boa introdução à lógica dos Testes de Hipóteses e a seus termos mais comuns. Trata-se de uma série de conceitos que precisam ser entendidos para dar sentido aos Testes de Hipóteses.

-  [O que são Testes de Hipóteses](#)
-  [Tipos de Erros e Nível de Significância](#)
-  [P-Valor](#)

Abaixo você encontra alguns materiais escritos sobre os Testes de Hipóteses.

-  [Teste de Hipóteses para uma e duas médias](#)
 -  [Testes de Hipóteses: Teoria](#)
 -  [Testes de Hipóteses: teoria e exemplos](#)
 -  [The Value Of P Values](#)
 -  [O que realmente significa o valor-p?](#)
-
-  [Exemplo de uso da Tabela T](#)
 -  [Exemplo de uso da Tabela Z](#)
 -  [Tabela Z e T](#)

Apêndice G

Situação e Material de Apoio para Aplicação Definitiva: Curso de Nutrição

TESTES DE HIPÓTESE PARA A MÉDIA POPULACIONAL

atividades
a distância

Nesta segunda incursão na PBL, os grupos vão analisar a seguinte situação:

Um pesquisador em nutrição quer avaliar o IMC de pacientes pós cirurgia bariátrica em sua cidade. Para isto levantou uma amostra destes pacientes gerando o arquivo "IMC pós bariátricos" encontrado abaixo. Use esta amostra para ajudar o pesquisador.

As perguntas a seguir ajudarão no processo de entendimento de um Teste de Hipótese para uma Média:

1. Qual é a média da amostra?
2. Considerando que um IMC 23 indica eutrofia, existe alguma chance (probabilidade) da população ter média 23, mesmo que a amostra não tenha média 23? Esta chance é grande ou pequena?
3. Se a população tiver média 23, existe alguma chance de se retirar uma amostra cuja média não é 23?
4. Se a população tiver média 23, qual a chance de tirar uma amostra com a média que obtivemos? Esta chance é grande ou pequena? O que significa grande ou pequeno neste contexto?
5. Conduza um teste de hipóteses para uma média com base na amostra levantada e decida se a média populacional é de 23:
 - a) Estabeleça a hipótese a ser testada e a alternativa. Escreva, com suas palavras, o que elas significam dentro do contexto desta situação.
 - b) Defina o nível de significância, justifique a sua escolha e explique o seu significado dentro do contexto da questão.
 - c) Defina a região de aceitação e rejeição de H_0 .
 - d) Calcule a estatística do teste e a localize no gráfico desenhado em c).
 - e) Calcule o valor de p adequado e explique o seu significado dentro do contexto da questão.
 - f) Decida sobre rejeitar ou não a hipótese nula e estruture uma resposta escrita para a situação.

Não esqueça que na PBL, o grupo analisa a situação, faz um levantamento do que já sabe e do que é necessário aprender (questões de aprendizagem), distribui as questões de aprendizagem para que os integrantes possam pesquisá-las e compartilharem o que descobriram no próximo encontro, procurando solucionar a situação.



IMC pós bariátricos

MATERIAL DE APOIO

O material a seguir ajudará o grupo a conduzir os Testes de Hipóteses apropriados para resolver as questões da situação.

A sequência de vídeos abaixo dá uma boa introdução à lógica dos Testes de Hipóteses e a seus termos mais comuns. Trata-se de uma série de conceitos que precisam ser entendidos para dar sentido aos Testes de Hipóteses.

-  [O que são Testes de Hipóteses](#)
-  [Tipos de Erros e Nível de Significância](#)
-  [P-Valor](#)

Abaixo você encontra alguns materiais escritos sobre os Testes de Hipóteses.

-  [Teste de Hipóteses para uma e duas médias](#)
-  [Testes de Hipóteses: Teoria](#)
-  [Testes de Hipóteses: teoria e exemplos](#)
-  [The Value Of P Values](#)
-  [O que realmente significa o valor-p?](#)

-  [Exemplo de uso da Tabela Z](#)
-  [Exemplo de uso da Tabela T](#)
-  [Tabela Z e T](#)

Apêndice H

Outras Situações Envolvendo os THMP

Situação 1

ESCALA DE AUTOEFICÁCIA

A Escala de Autoeficácia Geral (EAG) tem por objetivo mensurar o quanto o sujeito acredita em sua própria capacidade para alcançar os resultados por ele desejados.

Levantamento dos escores

A Escala de Autoeficácia Geral é constituída por 14 itens positivos: 1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 14, 15, 17 e 19. São itens do tipo "Eu sou capaz de superar muitos desafios".

Também, fazem parte da escala oito itens negativos: 8, 11, 13, 16, 18 e 20. São itens como "Quando problemas inesperados acontecem eu não lido bem com eles".

O escore bruto do sujeito é obtido pela soma das respostas dadas a cada item.

Entretanto, os itens negativos (e apenas eles) devem ter seu escore invertido antes que se faça a soma do escore total do sujeito. Vamos considerar um exemplo. Um sujeito marcou os seguintes números nos itens da escala:

Item 1) 5	*Item 11) 2
Item 2) 5	Item 12) 4
Item 3) 4	*Item 13) 4
Item 4) 5	Item 14) 4
Item 5) 4	Item 15) 5
Item 6) 5	*Item 16) 2
Item 7) 5	Item 17) 5
*Item 8) 1	*Item 18) 1
Item 9) 3	Item 19) 3
Item 10) 4	*Item 20) 1
Obs: itens negativos estão marcados com asterisco	

Figura H.1: Exemplo de marcação de respostas na Escala de Autoeficácia

Para os itens negativos, podemos utilizar a seguinte tabela:

Escore marcado pelo sujeito	Escore que deve ser somado
1	5
2	4
3	3
4	2
5	1

Figura H.2: Tabela de conversão de itens negativos na Escala de Autoeficácia

Para o exemplo dado, o escore bruto do sujeito será $5 + 5 + 4 + 5 + 4 + 5 + 5 + *5 + 3 + 4 + *4 + 4 + *2 + 4 + 5 + *4 + 5 + *5 + 3 + *5$ (os itens que tiveram o escore invertido estão marcados com asterisco). A soma desses escores foi 86. Para entendermos se esse valor é alto, baixo ou médio (ou seja, para avaliar qual é a posição do sujeito com relação à população), devemos utilizar a tabela de normas. O escore bruto do sujeito é 86 (coluna do meio na tabela) e isso corresponde ao percentil 90 (coluna da esquerda). Isso significa que 90% da população têm escore bruto de autoeficácia menor que o dele. Podemos concluir que esse sujeito tem alta autoeficácia

geral e, por isso, provavelmente é um sujeito que se sente capaz de vencer desafios, que acredita que pode atingir os objetivos desejados, que persiste em seus planos, pois acredita em sua capacidade de conseguir o que quer.

Mulheres			Homens		
Percentil	Escore bruto	Escore T	Percentil	Escore bruto	Escore T
5	53	33	5	59	32
10	59	37	10	62	35
15	61	40	15	67	41
20	62	41	20	68	42
25	64	43	25	70	44
30	67	45	30	71	45
35	68	46	35	72	46
40	70	48	40	74	48
45	71	49	45	75	49
50	72	50	50	77	51
55	74	52	55	78	52
60	76	53	60	79	53
65	77	55	65	81	55
70	79	56	70	81	55
75	81	58	75	82	56
80	82	59	80	84	58
85	84	61	85	86	60
90	87	64	90	89	64
95	90	66	95	91	66
M 72,2			M 75,9		
DP 11,1			DP 9,6		

Figura H.3: Tabela de normas para interpretação de resultados na Escala de Autoeficácia

Defina um valor para o escore da EAG que você considera um resultado a partir do qual o sujeito seja considerado "eficaz". Tome os dados coletados em "amostra Autoeficácia alunos UFN"¹ e teste a hipótese que afirmar que os alunos da UFN podem se considerar eficazes.

A situação 1 foi adaptada de: HUTZ, Claudio Simon. Avaliação em psicologia positiva.

Artes Médicas Editora, 2014.

Situação 2

¹Estes dados foram disponibilizado aos alunos como planilha do Excel.

Cem pacientes foram estudados quanto à taxa de fenilalanina no soro. A média obtida foi 1,3 mg/100mL e supôs-se de que tais indivíduos seriam uma amostra casual da população de indivíduos normais, onde a média é 1,4 e o desvio padrão, 0,32 mg/100mL. Considerando-se um nível de significância de 0,01, os dados estão de acordo com a suposição feita?

Situação 3

Em 2012 um estudo sobre o Número de Suicídios (NS) publicado pela Organização Mundial da Saúde (OMS) afirma que ocorrem anualmente, em média, 14,5 suicídios para cada 100.000 pessoas em todo mundo. A partir de uma amostra selecionada ao acaso com 24 países que apresenta um NS médio de 13 com DP de 9 suicídios para cada 100.000 pessoas, teste a afirmação da OMS.

Situação 4

Um pesquisador afirma que o escore médio de alexitimia avaliado pela Escala de Toronto (EAT20), para os alunos de uma dada faculdade, é menor que 30 pontos. Uma amostra aleatória com $n = 10$ estudantes da faculdade acusa escore médio de 25 pontos com desvio padrão de 5 pontos. Teste a hipótese do pesquisador considerando um nível de significância de 10%.

Situação 5

Considere a seguinte amostra: 3, 3, 2, 1, 0, 6, 5 e 4. Use estes dados para testar a hipótese nula $H_0 : \mu = 5$ contra a alternativa $H_1 : \mu < 5$ no nível de significância 0,05.

Apêndice I

Instrumento para Análise do Rendimento

Este instrumento foi aplicado por meio da ferramenta "Questionário", do Moodle, no laboratório de informática, com as alternativas mostradas de forma aleatória aos alunos. A seguir, é mostrada uma das possibilidades de visualização do instrumento, pelos alunos. As alternativas corretas estão em negrito.

1. Um fabricante de sabão líquido utilizado para fins hospitalares disse que, em média, o conteúdo de cada embalagem comercializada é de 10 litros. O responsável pela seção de compras de um hospital selecionou aleatoriamente sete embalagens de um lote, encontrando os seguintes conteúdos, em litros:

10,2 9,8 10,4 9,8 10 10,2 9,6

Sendo μ o conteúdo médio de sabão de todo o lote (a população), testar a hipótese nula $H_0 : \mu = 10$ litros contra $H_1 : \mu \neq 10$ litros, considerando uma chance de erro tipo I de 0,05. De acordo com o resultado do teste, o fabricante está dizendo a verdade quando afirma que o conteúdo de cada embalagem comercializada é de

10 litros, em média?

Escolha uma:

- (a) **Sim, pois p é maior do que o nível de significância**
- (b) Não, pois p é maior do que o nível de significância
- (c) Sim, pois p é menor do que o nível de significância
- (d) Não, pois p é menor do que o nível de significância
- (e) Não, pois p é diferente do nível de significância

A próximas questões referem-se ao problema a seguir:

Foi medida a pressão diastólica de 32 crianças com idade entre 5 e 6 anos, escolhidas aleatoriamente em uma determinada comunidade. A média da pressão diastólica obtida com base nessa amostra foi de 56 mmHg, com desvio padrão de 17,9 mmHg. Sabendo que os valores obtidos nacionalmente dizem que a média da pressão diastólica é de 64 mmHg para as crianças com idade entre 5 e 6 anos. Procurando testar se esta amostra de 32 crianças provém de uma população com média da pressão diastólica maior do que a média nacional, considerando um nível de significância de 5%, responda às seguintes questões:

2. As hipóteses, nula (H_0) e alternativa (H_1) do teste, são:

Escolha uma:

- (a) $H_0 : \mu = 56$ mmHg; $H_1 : \mu \neq 56$ mmHg
- (b) $H_0 : \mu = 64$ mmHg; $H_1 : \mu \neq 64$ mmHg
- (c) $H_0 : \mu = 56$ mmHg; $H_1 : \mu > 56$ mmHg
- (d) $H_0 : \mu = 64$ mmHg; $H_1 : \mu < 64$ mmHg
- (e) **$H_0 : \mu = 64$ mmHg; $H_1 : \mu > 64$ mmHg**

3. O teste apropriado é um

Escolha uma:

- (a) Teste z ou t, pois a amostra é próxima de 30
- (b) Teste t, pois a amostra é maior do que 30
- (c) Teste bilateral, pois a amostra é maior do que 30
- (d) Teste unilateral, pois a amostra é maior do que 30
- (e) **z, pois a amostra é maior do que 30**

4. O valor da estatística do teste é de

Escolha uma:

- (a) $z = - 10,69$
- (b) $t = 10,69$
- (c) $t = - 2,53$
- (d) $t = - 10,69$
- (e) **$z = - 2,53$**

5. O p-valor calculado para o teste é de

Escolha uma:

- (a) **$p = 0,0057$**
- (b) $p = 0,0011$
- (c) $p = 0,071$
- (d) $p = 0,057$
- (e) $p = 0,0114$

6. Sobre a tomada de decisão do teste

Escolha uma:

- (a) Aceita-se H_0 pois $p > \alpha$
- (b) O teste foi inconclusivo
- (c) **Rejeita-se H_0 pois $p < \alpha$**
- (d) Não rejeita-se H_0 pois $p < \alpha$
- (e) Rejeita-se H_0 pois $p > \alpha$

7. Relativamente ao nível de significância, é correto afirmar que ele

Escolha uma:

- (a) Deve ser maior do que o p-valor
- (b) **Deve ser escolhido antes do teste**
- (c) Deve ser escolhido de acordo com o p-valor
- (d) Deve ser menor do que o p-valor
- (e) Deve ser escolhido depois do teste

8. Relativamente ao p-valor, é correto afirmar que ele

Escolha uma:

- (a) É a probabilidade de H_0 ser falsa
- (b) É a probabilidade de H_0 ser verdadeira
- (c) **É calculado considerando-se H_0 verdadeira**
- (d) Deve ser sempre menor do que α (alfa)
- (e) Deve ser sempre maior do que α (alfa)

9. Relativamente a um teste de hipótese para a média populacional. A afirmação verdadeira é

Escolha uma:

- (a) O teste não pode ser conduzido quando a população não é normalmente distribuída
- (b) Quando se rejeita a hipótese alternativa, automaticamente se aceita a hipótese nula
- (c) O teste comprova se a hipótese nula é verdadeira ou se ela é falsa
- (d) O teste verifica se a média amostral é igual à média populacional
- (e) **Para o cálculo do p-valor, o teste pressupõe que a hipótese nula é verdadeira**

10. Um novo medicamento promete reduzir a pressão arterial sistólica (pas) de pacientes hipertensos. Foi selecionada uma amostra aleatória de pacientes hipertensos e conduzido um teste de hipóteses para a média populacional que rejeitou a hipótese nula de igualdade entre a média populacional e o valor de referência para a média da pas de hipertensos. A afirmação verdadeira é

Escolha uma:

- (a) **O medicamento provocou uma queda significativa da pas de hipertensos**
- (b) Por não ter sido mencionado, o nível de significância adotado foi de 5%
- (c) A probabilidade de erro do tipo 1 foi menor do que o nível de significância
- (d) O poder do teste foi alto, pois a hipótese nula foi rejeitada
- (e) O novo medicamento reduziu a pas de todos os pacientes da amostra

Apêndice J

Termo de Consentimento

Termo de consentimento

Prezado(a) acadêmico dos cursos da área da saúde da Universidade Franciscana – UFN

Sou doutorando em Educação na Universidade de Burgos – Espanha. Minha pesquisa visa contribuir para a melhoria do ensino de estatística, em particular do conteúdo de testes de hipóteses, de estudantes da área da saúde.

Os estudantes que optarem por participar deste estudo, devem estar matriculados na disciplina de bioestatística.

Seu consentimento é voluntário e envolverá a participação nas atividades didáticas e levantamento de dados subsequentes, sem comprometimento da nota da disciplina.

A identidade dos participantes não será revelada em nenhuma fase da pesquisa.

Conto com a sua participação e saliento a importância do desenvolvimento de uma cultura educacional que proporcione espaços de discussão e construção do conhecimento.

Subscrevo-me, atenciosamente,
Rodrigo Fioravanti Pereira

Eu, _____, aluna (o) do curso de _____ concordo em participar das atividades didáticas referentes à pesquisa intitulada TESTES DE HIPÓTESES: ANÁLISE DO DOMÍNIO DESTE CAMPO CONCEITUAL POR ALUNOS DA ÁREA DA SAÚDE, conduzida pelo professor Rodrigo Fioravanti Pereira, sem prejuízo de minha nota na disciplina de bioestatística, estando ciente do objetivo do estudo e de meu envolvimento. Concordo que minha produção intelectual seja divulgada nos meios científicos, respeitando meu anonimato.

Assinatura do participante

Santa Maria, _____ de _____ de 2020.